Wissenschaftszentrum Weihenstephan
für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Produktions und Ressourcen Ökonomie

# Culture and Rural Development

## An Empirical Analysis

David Johannes Wüpper

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Agrarwissenschaften genehmigten Dissertation.

Vorsitzender: Prof. Dr. J. Kollmann

Prüfer der Dissertation

1. Prof. Dr. J. Sauer
2. Prof. Dr. G. Buchenrieder
3. Prof. Dr. J. Kantelhardt
   Universität für Bodenkultur Wien / Österreich

Die Dissertation wurde am 13.6.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 12.9.2016 angenommen.

# Content

# Zusammenfassung

Geschichte erklärt ökonomische Unterschiede – zwischen Weltregionen, Ländern, Regionen und Individuen. Ein Grund dafür ist Kultur. Kultur ermöglicht es uns, Lernkosten einzusparen, weil Verhaltensmuster unser Vorfahren übernommen werden, ohne das wir sie genau verstehen müssen. Anthropologen sind sich weitgehend einig, dass es unsere Fähigkeit der Imitation ist, die uns Menschen so anpassungsfähig macht, im Vergleich zu anderen Spezies, die entweder weniger oder einfach schlechter imitieren. Allerdings macht uns Kultur auch weniger anpassungsfähig als es häufig in der Ökonomie für den Homo Oeconomicus angenommen wird. Schließlich führt Kultur dazu, dass Verhaltensmuster oft einfach imitiert werden, ohne in Frage gestellt zu werden.

In dieser Doktorarbeit sind mehrere Kapitel dem Konzept der wahrgenommenen Selbst-Wirksamkeit gewidmet. Die Bedeutung der Selbst-Wirksamkeit wurde vor Allem von Psychologe Albert Bandura erforscht. Es beschreibt wie sehr eine Person daran glaubt, die Fähigkeit zu haben, ihre selbst gewählten Ziele zu erreichen. Dieser Glaube beeinflusst, welche Ziele gewählt werden, und wie effektiv ihre Erreichung verfolgt wird. Personen mit niedriger Selbst-Wirksamkeit in einer Domäne vermeiden sie entweder vollständig, oder sind unmotiviert genügend zu investieren – insbesondere wenn auch noch weitere Hürden hinzukommen.

Eine der Haupt-Forschungshypothesen dieser Arbeit war, dass wahrgenommene Selbst-Wirksamkeit historisch-kulturell bedingt ist und Relevanz für die Entwicklungsökonomie hat.

Die Arbeit beginnt mit einer Analyse, ob koloniale Erfahrungen in Ghana den Erfolg von Produktions-Verträgen in der Landwirtschaft beeinflussen.

Zu Kolonialzeiten etablierte die britische Regierung Kooperativen für den Kakao Export und christliche Missionare etablierten Schulen. Der damalige Erfolg der Kooperativen beeinflusst noch heute die Selbst-Wirksamkeit der Landwirte im Bezug auf globale Wertschöpfungsketten, und die christlichen Missions-Schulen beeinflussen noch immer ihr

Sozial-Kapital. Beide Variablen sind sehr wichtig für den Erfolg der Vertragslandwirtschaft, welche wiederum ein wichtiges Werkzeug zur Armutsbekämpfung ist.

Sogar noch früher als die kolonialen Erfahrungen sind Erfahrungen mit vorindustriellen Produktionssystemen. In Regionen, in denen die ökologischen Voraussetzungen den Anbau von Getreide begünstigten, entwickelten die Landwirte hohe Investitions-Selbst-Wirksamkeit, weil Getreide Investitionen belohnte. In anderen Regionen, in denen die Biogeographie eher andere Anbausysteme bevorzugte, entwickelten die Landwirte eher niedrige Selbst-Wirksamkeit, weil zum Beispiel Wurzeln und Knollen wie Cassava und Yams Investitionen weniger erforderten und auch weniger belohnten. Im heutigen Ananas-Anbau spielen Investitionen eine sehr wichtige Rolle. Interessanterweise investieren die Nachfahren von getreideanbauenden Landwirten deutlich mehr als die Nachfahren von anderen Landwirten, weshalb sie deutlich höhere Einkommen haben.

Eine besondere Eigenschaft der Selbst-Wirksamkeit ist, dass sie die Reaktion auf Rückschläge beeinflusst. Individuen mit hoher Selbst-Wirksamkeit reagieren mit erhöhter Motivation, während Individuen mit niedriger Selbst-Wirksamkeit möglicherweise ganz aufgeben. Für die Landwirte in Ghana ist der Regen eine wichtige Einkommens-Determinante. Spannenderweise kann man in der Tat beobachten, dass Landwirte mit hoher Selbst-Wirksamkeit auf Dürren mit der Übernahme wassersparender Innovation reagieren, während Landwirte mit niedriger Selbst-Effektivität sich gar nicht anpassen.

Diese Ergebnisse führen natürlich zu der Frage, welche Faktoren es wohl Personen und Regionen ermöglichen, bessere ökonomische Ergebnisse zu erzielen, als von ihrer Geschichte prognostiziert. Die Antwort: Bildung und Sozial-Kapital.

Training in ausgewählten Innovationen könnten ebenfalls helfen. Es ist allerding klar, dass Training nicht für alle Technologien gleich effektiv ist. Im Hinblick auf nachhaltige Intensivierungs-Technologien ist das Ergebnis, dass eher simple Innovationen leicht von anderen Landwirten gelernt werden können, wodurch deutlich weniger Training notwendig ist, als für komplexere Innovationen, die stark und lange vom Training profitieren.

Zum Ende wird global eine ganz andere Beziehung zwischen Ökonomie und Kultur untersucht. Die Weltkulturerbeliste der UNESCO soll besondere Orte beschützen und kommunizieren. Eine wichtige Frage für die UNESCO ist jedoch, warum sich nicht alle gelisteten Orte klar als Weltkulturerbestätte identifizieren. Die Antwort ist eine Reihe orts- und regions-spezifischer Anreize, häufig verbunden mit Tourismus-Einnahmen. Die Kosten der Weltkulturerbe-Vermarktung und die Motivation das Programm voranzubringen spielen im Gegensatz kaum eine Rolle. Ein großer Anteil des Verhaltens ist rein kulturell bedingt, sodass Orte im Nahen Osten zum Beispiel gar nicht als Weltkulturerbe-Stätte vermarktet werden, und besonders in Asien sehr stark.

# Summary

History is an important determinant of current economic development. One reason is cultural learning, which includes imitating behaviors from ancestors in order to save individual learning costs. Amongst anthropologists, there is widespread agreement that it is cultural learning that makes humans so adaptive in comparison to other species, which imitate less or worse. Nevertheless, culture also makes humans less adaptive than economists assume for the homo economicus because humans imitate many behaviors without appraisal, inefficient behaviors might persist for a long time before they are changed.

In this PhD research, much attention is focused on a cultural trait called self-efficacy. The concept has been developed by psychologist Albert Bandura and describes how much a person believes to have the ability to achieve self-chosen goals. Research has shown that self-efficacy affects which goals are chosen and how effectively they are pursued. Individuals with low self-efficacy in a domain either avoid it, or are unmotivated to invest sufficient effort, especially in the face of obstacles.

The thesis begins with an investigation of whether colonial experiences persist to affect current contract farming performance in Ghana. During colonial times, the British government established cocoa export cooperatives and Christian missionaries established schools. The performance of the cocoa export cooperatives is found to have shaped the long-term self-efficacy of the farmers in regard to the profitability of such global value chains and the Christian missionary schools persistently lowered village level social capital. Thus, historically rooted cultural differences currently explain the performance of contract farming in different communities.

Even earlier causes of divergent cultures are the experiences with pre-industrial subsistence farming systems. Where the ecological setting incentivized cereal farming, farmers were rewarded for agricultural investments and thus developed self-efficacy regarding agricultural investments. Where the ecology incentivized other farming systems based on roots, tubers, or tree crops, investments were less rewarding and farmers developed lower investment self-

efficacy. These differences are found to significantly explain income differences amongst Ghana's current pineapple farmers. The causal channel are investments, which are critical for the profitability of pineapple and which are determined by the farmers' investment self-efficacy.

A special feature of self-efficacy is furthermore, how people react to adversity. Whereas high self-efficacy leads people to increase their efforts after failure, low self-efficacy leads to decreased efforts. It is found that farmers with high self-efficacy are able to mitigate a significant share of lost income from droughts. The reason is that they are more likely to adopt a climate smart innovation that conserves water when rainfall decreases. Their peers with low self-efficacy are not found to adapt.

Investing which farmers achieve higher incomes than predicted by ancestral' experiences, it is the well-known variables education and social capital. Thus, overcoming history is not found to require special policies, at least for the pineapple farmers in Ghana.

Agricultural trainings about innovations are also a potential policy tool to increase rural incomes in Ghana. However, a significant effect is only found for more complex innovations, whereas simpler innovations can easily be learned from other farmers.

Globally, a very different relationship between culture and economics is investigated. Attempting to explain why not all World Heritage sites are promoted as such, it is found that site and region specific, economic variables explain the pattern well – whereas constraints and the collective benefit do not matter much. To strengthen the brand, it is thus either necessary to help more sites to benefit, or to make promotion mandatory.

# INTRODUCTION: CULTURE AND RURAL DEVELOPMENT

It is widely acknowledged that economic development is the outcome of history (Acemoglu and Robinson 2012, Nunn 2013, Spolaore and Wacziarg 2013). Now, attention has turned to the specific mechanisms behind this. Especially the role of culture is receiving widespread interest - one the one hand, because it is so fundamental to human behavior, and on the other, because it is so difficult to quantify.

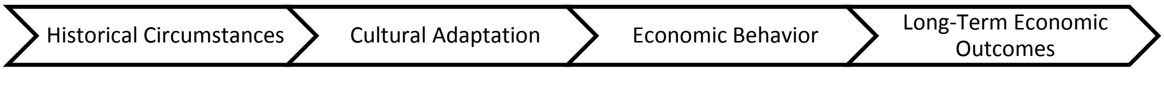
As Nunn (2012) argues, “different societies make systematically different decisions when faces with the same decision with exactly the same available actions and same payoffs. A natural interpretation of these systematic differences is that specific decision-making heuristics evolved in different societies due to the particular environments or histories of the groups”. This is culture.

The reason why human decision making is cultural (and not “rational” as often assumed in economic models for simplicity) is that our environment is too complex to make rational decisions (Simon 1982, Henrich et al. 2001, Gigerenzer and Gaissmaier 2011). The research of Damasio et al. (1994) and Bechara et al. (1997) show how our ability to make decisions crucially depends on our emotions and feelings, so that we often feel information before we know it. This is “better than rational” (Cosmides and Tooby 1994). Humans could never have conquered all major habitats across the globe with purely individual and rational learning, according to Boyd et al. (2011). Instead, humans learn from each other – horizontally, e.g. from neighbors and vertically, e.g. from parents. What is learned are heuristics, or simplified decision rules, which allow to save learning costs and enable roughly adapted behavior. Obviously, roughly adapted behavior is inferior to perfectly adapted behavior, but evolutionary, the costs of learning perfect behaviors were prohibitively high (Boyd and Richerson 1985, Richerson and Boyd 2008).

To give a few concrete examples, Nunn and Wantchekon (2011) find that Africa’s major slave trades persistently eroded inter-personal trust. During the slave trades, people learned not to trust others, because being too trusting often resulted in being exported as a slave. Today, trust is a major determinant of economic development (Knack and Keefer 1997), but because it is costly to learn who to trust, how much, and when, most people are generally more or less trustworthy towards different categories of people (e.g. family, friends, strangers) and they imitate largely the behavior of their social peers. Thus, even though higher levels of trust would result in higher economic growth, countries that were strongly impacted by the slave trades find it often difficult to develop such trust. A second example comes from Europe. Tabellini (2010) finds that historically better institutions (between 1600 and 1850) positively and persistently affected trust and respect for others, as well as confidence in the benefit of individual effort (also called self-efficacy – but more about this trait below). These cultural traits are found to explain the income differences within European countries.

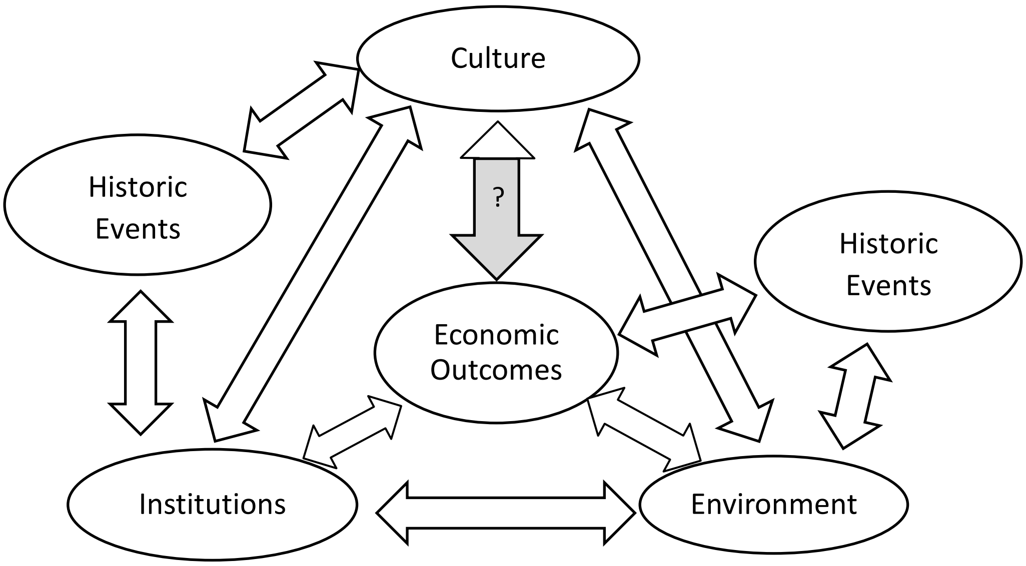
The two given examples show how culture is usually well adapted to long-term contexts and only slowly adjusts to new contexts. According to the Cultural Evolution theory of Boyd and Richerson (1985), Richerson and Boyd (2008), and Boyd et al. (2011), a popular strategy that has evolved throughout history is that younger generations imitate the behavior of older generations and if there is an incentive to do so, test different behavioral rules, and depending on a comparison, choose the one that is better. However, if learning signals are not clear, or if behavioral change has to be performed collectively (think of trust, which is only mutually beneficial when it is shared), culture can be very stable across time and different culture can co-exist in close geographic vicinity (Grosjean 2014).

For agricultural economists, incorporating culture into theoretical and empirical models is a logic step. It is common practice to include horizontal network effects in models explaining technology adoption and productivity (Sauer and Zilberman 2012, Maertens and Barrett 2013, Magnan et al. 2015). The next step is to take vertical network effects serious. The basic idea looks as follows:



Recently, especially work in development economics has attempted to increase the behavioral realism of its models by incorporating insights from psychology (Bertrand et al. 2004, Mullainathan 2005, Banerjee and Mullainathan 2008, Duflo et al. 2009, Banerjee and Duflo 2011, Shah et al. 2012, Mani et al. 2013, Mullainathan and Shafir 2013, Datta and Mullainathan 2014, Hanna et al. 2014). In the following chapters, a major contribution is to show that cultural evolution explains the observed behavioral phenomena (Morgan et al. 2015, Mesoudi 2016).

**Figure 1. Long-Term Impacts on Economic Outcomes**

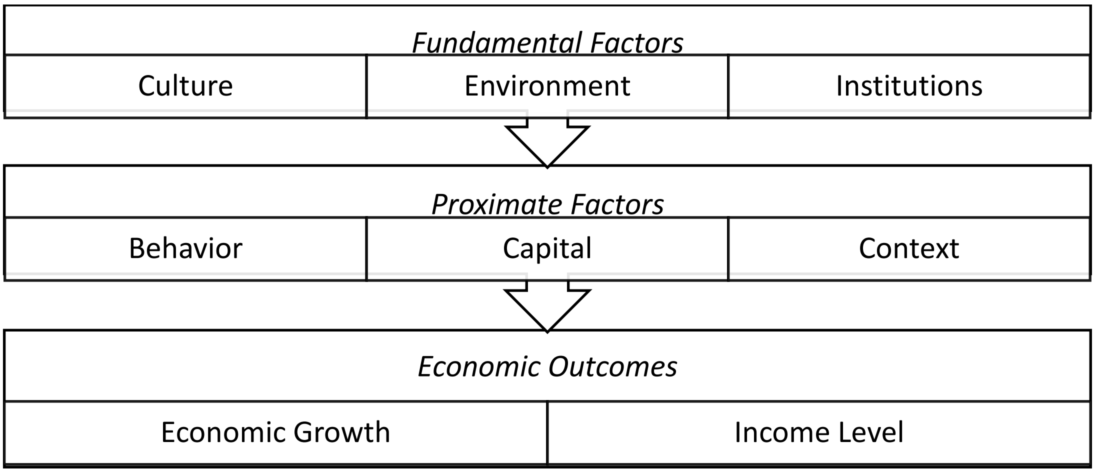



Notes: Figure 1 shows the complex interactions between the fundamental determinants of economic outcomes. Research about the effect of culture on economic outcomes thus must isolate the channel from culture to outcomes from the other channels.

The challenge of quantifying the economic effect of culture can be seen in figure 1, which shows the fundamental identification problem. We are interested in the effect of culture on economic outcomes, e.g. the economic pay-off from having inherited more patience or entrepreneurial spirit (because these were evolutionary advantageous to one's ancestors). The first thing that can be seen in figure 1 is that culture is endogenous because economic outcomes are not only affected by culture, but culture is also affected by economic outcomes (Inglehart and Baker

2000). Furthermore, the historic circumstances that shaped one's culture can also have an effect through other channels (institutions, capital accumulation). And finally, one's geography can have a persistent effect on one's behavior and economic outcomes – directly and independent from any effects on culture and institutions (Gallup et al. 1999, Sachs 2003). Thus, it is necessary to use advanced econometric techniques, to disentangle the causal effect of culture from all the other potential effects, such as the ones displayed in figure 1.

**Figure 2. Fundamental and Proximate Causes of Economic Outcomes**




Notes. Figure 2 shows different levels of determinants of economic outcomes. The fundamental factors could also be further divided into first order and second order fundamental factors, because both culture and institutions are shaped by the environment.

In this context, it is important to note the nature of culture as a *fundamental* factor for economic development, in contrast to *proximate* ones – as displayed in figure 2. This means that cultural traits do not directly compete with other incentives and constraints, but often, they explain them. As an example, a culturally inherited entrepreneurial spirit might explain why an economic agent got a credit, invested in a risky business, and increased her capital base. In such a case, one would not want to control for these outcomes when investigating the causal effect of entrepreneurial spirit on one's income, as it would control away much of the actual effect. On the other hand, it is clear that we must control for non-cultural variables, as a person with e.g. a higher capital endowment will usually find it easier to exploit business opportunities, independent from culture. Thus, to identify the causal effect of cultural traits, non-cultural influences must be controlled for, without controlling away the cultural influences.

The benefit of including culture in economic investigations is often large. First of all, it has often been argued that policy-failures can be attributed to a misinterpretation or deliberate ignorance of the explanation why current constraints exist in the first place (Harrison and Huntington 2000, Rogers 2010, Banerjee and Duflo 2011). Secondly, culture can sometimes explain behaviors that would seem strange without taking culture into account, such as a reluctance to adopt certain innovations (Rogers 2010), non-adoption to climate change because of a cultural norm to support one's kin in emergency situations (Di Falco and Bulte 2011), or the choice of different agricultural production systems independent from economic incentives but dependent on the background of the farmers' ancestral background (Richerson and Boyd 2008).

In most of the following chapters, the focus lies on the diffusion of innovations, such as a new value chain, new agricultural practices, or the adaptation to climate change. The main explanation is self-efficacy, which captures how much a person is convinced to have what it takes to achieve a chosen goal. This belief is a fundamental determinant of behavior because it affects what goals people set for themselves, how hard they try to achieve them, and how much adversity they can withstand before they give up (Bandura 1977, 1995, Maddux 1995, Bandura 1997, 2012). Self-efficacy affects human decision making at all levels because people only choose actions which they believe to be worthwhile. Only few people deliberately choose an unachievable goal and similarly few people deliberately invest much effort into actions that will have disappointing results.

Of course, most important choices in life are ambiguous and it is not clear whether a given goal can be achieved, whether one's performance will be satisfactory, or whether investments will pay off. Thus, a person's expectation might be the most important determinant for her choices and many observed behavioral puzzles can be understood once the beliefs of a person are known.

The diffusion of innovation is a good example. The idea that individuals have different adoption thresholds (so that those individuals with the lowest threshold adopt the innovation first) is well understood theoretically and consistent with the empirical data (Feder et al. 1985, Feder and Umali 1993, Zilberman et al. 2012). The question then is which individual differences determine the adoption thresholds. To understand the effect of self-efficacy it helps to consider a simple theoretical model.

A feasible foundation for a self-efficacy model is the target input model, which is often used to conceptualize learning processes about innovations (Foster and Rosenzweig 1995, Bandiera and Rasul 2006, Conley and Udry 2010). A short discussion of the model is provided by Bardhan and Udry (1999). By substituting inputs such as fertilizer or a new variety with self-efficacy, we can explore how self-efficacy develops in time and how it affects economic outcomes.

Suppose output $q_{it}$ is determined by the squared difference of the chosen technology usage $\alpha_{it}$ and the optimal technology usage $\beta_{it}$. This notion is general enough to accommodate both discrete and continuous technologies. The production function might look as follow:

$$q_{it} = 1 - (\alpha_{it} - \beta_{it})^2 \tag{1}$$

The optimal technology usage fluctuates around a mean due to independent and identically distributed (i.i.d.) shocks with mean zero and variance $\sigma_u^2$:

$$\beta_{it} = \beta_i^* + \mu_{it} \tag{2}$$

Farmer $i$ does not know $\beta_i^*$ at time $t$ but has beliefs about it, which are distributed $N(\beta_{it}^*/\theta_{it}, \sigma_{\beta_{it}}^2)$. In contrast to the standard target input model, the beliefs of some farmers are systematically biased in our model. $\theta_{it} > 1$ captures how much. If $\theta_{it} = 1$, the farmer is unbiased, for $\theta > 1$ the farmer does not belief to have the ability to achieve a sufficient performance level with the technology and thus discounts its profitability.

In time, farmers update their beliefs about $\beta_i^*/\theta_i$ but in contrast to the standard target input model, belief updating must not increase productivity. The intuitive explanation is that outcomes, learning and beliefs are all interdependent, so that it is possible to learn something wrong, which implies that belief updating can increase, decrease, or stabilize productivity.

Making the simplifying assumption of a costless input, we can interpret output as profit. Expected profit maximization then implies the following input choice:

$$\alpha_{it} = E_t(\beta_{it}) = \beta_i^*/\theta_i \tag{3}$$

and the following production function:

$$E_t(q_{it}) = 1 - \theta_{it} - \sigma_{\beta_{it}}^2 - \sigma_u^2 \tag{4}$$

which means that expected profit rises when the farmers learn about the true value of $\beta_i^*$ (as $\sigma_{\beta_{it}}^2$ and $\theta_{it}$ decline) and the more predictable their operating environment is (as $\sigma_u^2$ declines).

We might begin with learning by doing, for now without social learning.

Let us define $$\theta_{it} + \sigma_{\beta_{it}}^2 \equiv \tau_{it}$$

as total technology bias of the farmers, which has a systematic part (different degrees of self-efficacy) and a random part (uncertainty due to a lack of information). Furthermore,

$$\frac{1}{\theta_{it}+\sigma_{\beta_{i0}}^2} \equiv \rho_1$$

which denotes the accuracy and precision of a farmer's initial beliefs and

$$\frac{1}{\theta_{it}+\sigma_u^2} \equiv \rho_2$$

which denotes the accuracy and precision of a farmer's learning signals.

Farmers choose technology inputs, observe outcomes, and update their beliefs about $\beta_i^*$ using Bayes' rule, so that learning is a function of initial beliefs, learning signals, and the extend of trials:

$$\tau_{it} = \frac{1}{\rho_1 + I_{it-1}\rho_2} \tag{5}$$

where $I_{it-1}$ is an index summarizing both the number of trials and the effort put into each one.

This index is a direct function of the self-efficacy of the farmers:

$$I_{it} = f(\theta_{it}), \quad \text{so that} \quad \frac{\partial I_{it}}{\partial \theta_{it}} > 0$$

Because self-efficacy (a) increases the effort that is invested into each trial and (b) increases the number of trials, despite likely set-backs and disappointments, self-efficacy affects expected profits:

$$Eq_{it}(\theta_{it}) = 1 - \frac{1}{\rho_1 + \theta_{it}\rho_2} - \sigma_u^2 \tag{6}$$

So that

$$\frac{\partial Eq_{it}(\theta_{it})}{\partial \theta_{it}} = \frac{\rho_2}{(\rho_1 + \theta_{it}\rho_2)^2} > 0 \tag{7}$$

As long as learning positively affects productivity, self-efficacy positively affects productivity.

Thus, farmers with high self-efficacy start off with a more accurate technology belief and their beliefs converge faster to the true value, whereas farmers with low self-efficacy start off with a less accurate technology belief and then might never converge to the true value.

Let us now consider social learning.

Self-efficacy has four sources: (a) persuasion, (b) observing the experiences of social peers, (c) own experiences, and (d) emotional cues. So a farmer might have a certain degree of self-efficacy because she was persuaded by someone to have or to lack certain abilities; she probably has inferred her abilities from observing the choices and outcomes of people she judges similar to herself; certainly, she also compared and updated her beliefs after experiencing the outcomes of her choices; and finally, her degree of self-efficacy might be shifted by her general emotions and personality. Important to note is that self-efficacy is a self-reinforcing belief (which can also be seen in the equations above), so a farmer's experiences are more likely to be positive if she had higher initial self-efficacy, and the same is true for the social peers of the farmer, so that entire social networks might learn to have distinct levels of ability, as a function of their initial self-efficacy.

In this section, we will see how social learning can make self-efficacy an even stronger reinforcer of initial self-efficacy beliefs than it was with pure individual learning.

Suppose farmer $i$ can observe the technology input choice of her neighbors $j$ – possibly not entirely correctly, because of some observational error $\sigma_{\varepsilon}^2$. This means, the farmer observes $\beta_{jt} + \epsilon_{jt}$ with $\epsilon_{jt} \sim N(0, \sigma_{\epsilon}^2)$, assuming that $\sigma_{\epsilon}^2$ is known.

We then define $$\frac{1}{\sigma_u^2+\sigma_\varepsilon^2+\theta_{jt}} \equiv \rho_3 \quad < \rho_2$$

Which is the accuracy and precision of network learning signals, which depends on the network context, how precisely choices and outcomes can be observed, and the peers' self-efficacy.

We thus have $$\tau_{it} = \frac{1}{\rho_1+I_{i,t-1}(\theta_{it})\rho_2+N_{j,t-1}(\theta_{jt})\rho_3} \tag{8}$$

Farmer $i$ now also learns from the trials of her peers, in addition to her own trials. Self-efficacy affects the accuracy of her own beliefs and those of her peers, and it also affects how fast she and her peers learn, so that some networks converge much quicker to the true value of $\beta_j^*$ than others. In addition, there is one more effect: The self-efficacy of her peers can affect the self-efficacy of farmer $i$ and vice versa: $$\tau_{it} = \frac{1}{\rho_1+I_{i,t-1}(\theta_{it}(\theta_{jt}))\rho_2+N_{j,t-1}(\theta_{jt}(\theta_{it}))\rho_3} \tag{9}$$

So that $$\frac{\partial E q_{it}(\theta_{it},\theta_{jt})}{\partial \theta_{jt}} = \frac{\rho_2}{(\rho_1+(\theta_{it}(\theta_{jt}))\rho_2+(\theta_{jt}(\theta_{it}))\rho_3)^2} > 0 \tag{10}$$

Thus, a farmer's level of self-efficacy affects her own productivity and that of her social peers which means that high self-efficacy produces a positive externality but low self-efficacy produces a negative one. This can lead to a self-limiting dynamic in communities in which low level pursuits are chosen without reappraisal of actual abilities and no farmer ever learns that a more profitable production would be possible under current circumstances.

The model suggests that self-efficacy can be understood as a Bayesian prior about the ability to profit from an innovation. However, it is not updated in standard Bayesian fashion because the prior directly affects what is subsequently experienced. Finally, what other learn can strongly affect what an individual learns, so that entire communities can be locked into a Pareto-suboptimal equilibrium.

Arguably, one reason why self-efficacy is usually missing from innovation diffusion models is the challenge of identifying its effect. It is perhaps plausible that a farmer will not adopt an innovation if she does not think she can increase her profit with it. It is however a challenge to disentangle the effect of this belief from unobserved performance determinants. Perhaps farmer with more experience or better education are both more likely to profit from the adoption of an innovation and to expect this. Other variables that might lower the adoption threshold and increase a farmer's self-efficacy include financial means, information and insurance networks, infrastructure, biogeography, and many more. Some of these variables are easily observable but others are not. Thus, either randomized control trials (RCT) or advanced econometric techniques (AET) are required to identify the causal effect of self-efficacy. An example of an RCT is Bernard et al. (2014). In their study, they manipulated the self-efficacy of Ethiopian smallholder farmers by showing the treatment group a documentary about successful businesses that were started by social peers of the farmers. The control group watched an "uninformative" TV show. Bernard et al. (2014) find a significant causal effect of self-efficacy on aspirations and especially educational investments, savings, and loan taking. The robustness of these results however comes at a price. Because self-efficacy is experimentally manipulated, little can be learned about the existing differences in self-efficacy and only short term effects of a higher degree of self-efficacy can be investigated. Why do Ethiopian smallholder farmers have low levels of self-efficacy? How stable are differences in self-efficacy? Do these differences have historic roots? What are the effects of long-term developed self-efficacy?

To address these questions AET are attractive. In the following, we thus rely heavily on instrumental variables, such as used by Acemoglu and Robinson (2001) who use the historic roots of current institutions to identify their causal effect on economic performance. We develop novel instruments that are historic and exogenize the cultural trait self-efficacy.

This improves our understanding of the causal chain from historic circumstances, over human adaptations to those circumstances, to the causal effects of such adaptations. Thus, we understand why we observe differences in self-efficacy, and how this affects economic outcomes.

Most of the research described in the following chapters is based on data collected amongst smallholder pineapple farmers in Ghana. The main reason is that Ghana's historic development is advantageous to the study of culture, because (a) within the country, there are many different ethnicities with different histories and cultures, which are now subject to the same national institutions and economic context, (b) the history of Ghana offers a rich set of interesting events, which have often been quantified, and (c) economic development is dynamic in Ghana, so there is much spatial and temporal variation to be explained.

In the first chapter, it is investigated how colonial experiences shaped self-efficacy regarding more formal value chains and also social capital. It is found that both cultural traits (one individual and one collective) matter a great deal for the performance of contract farming.

In the second chapter, it is analyzed how different historic farming systems led to different levels of self-efficacy regarding agricultural investments. It is found that this explains much of the current investment and income differences amongst the farmers.

In the third chapter, the hypothesis is tested, that farmers with high self-efficacy react different to the experience of decreased rainfall than farmers with low self-efficacy. It is found that farmers with high self-efficacy react by adopting a climate-smart technology, whereas farmers with low-self do not.

The next two chapters, it is investigated what policy makers need to do to support the farmers most effectively.

In the fourth chapter, the current farm incomes are predicted using two historical events: The experience of the trans-Atlantic slave trade and the main crop that was grown in one's ancestral society. The causal mechanism that connects these variables with current farm incomes is different. Whereas the slave trade eroded social capital, the dependence on different crops shifted investment self-efficacy. Thus, an interesting question is which factors explain why some farmers achieve higher incomes than previously predicted. Perhaps surprisingly, social capital and education are found to have the same, positive effect. This establishes that a wide range of historically rooted cultural differences can be mitigated by the same set of well-known variables.

Whether recent development programs in Ghana have successfully fostered the adoption of sustainable innovations is investigated in chapter five. As most initiatives have focused on the provision of agricultural training, two such trainings are investigated and compared. Such an analysis, again, poses a challenge for the identification of causal effect. Because the farmers might learn about technologies from trainings or from other farmers, both variables need to be identified. However, trainings might be selectively offered or taken up by farmers who are different from those who do not participate. In the worst case, farmers who already chose to adopt a technology seek out training to learn how to use it. And homogenous group behavior might look like peer-learning but it could also be individuals reacting to the same incentives and constraints or acting according to shared individual characteristics. The instrumental variables used to identify the effect of learning from trainings and from peers are the lagged values from adjacent communities. To use peer-learning as example, how many farmers in a community adopt an innovation has little to do with the incentives, constraints, or characteristics in an adjacent farmers' group. However, how many farmers in a farmers' group adopt an innovation is usually highly correlated with the adoption rate in adjacent farmers' groups because of spatial continuities. The required assumption is that peer networks are partially transitive, meaning that some farmers know each other across groups, but not all. The finding is that training and peer learning are strong substitutes, so that the extent of possible peer learning determines whether trainings are beneficial to the farmers or not.

The last chapter investigates a different aspect of culture. As described by Harari (2014), we do not only call inherited heuristics culture but we also call things that transfer such heuristics culture. Examples are theaters, architecture, or music, which all transport cultural messages. Thus culture shapes culture and how people respond to culture depends on their culture. To investigate this issue, the last chapter is concerned with the World Heritage Program of the United Nations Educational, Scientific and Cultural Organization (UNESCO). The core of the program is a list of outstanding locations all around the world that are deemed humanities cultural and natural heritage according to a list of criteria and expert evaluations.

How much the individual locations are marketed as World Heritage sites is an important issue for UNESCO, because its World Heritage brand is build around its outstanding locations, which promote it. Then, World Heritage brand equity can be used to support economic development (Arezki et al. 2009, Rebanks 2009, Ryan and Silvanto 2011, Licciardi and Amirtahmasebi 2012). However, the logic of collective action might create an incentive for those sites with the highest potential to not contribute, because e.g. already popular tourism locations do not gain as much as sites with only few visitors, that try to use World Heritage as promotion argument. If the site managers are rational, their World Heritage promotion is solely bases on a cost benefit analysis. Alternatively, their behavior might be other-regarding (social preferences) and culturally biased (e.g. valuing World Heritage above or below its economic value). Chapter six thus analyzes a global sample of World Heritage sites in a big-data spatial econometrics framework and finds that economic incentives and culture are the main explanations.

Table 1 on the next page gives an overview of the following chapters and their contribution.

Table 1. Contribution in the following chapters.

| | Topic | Question | Hypothesis | What is new |
|---|---|---|---|---|
| 1 | Contract Farming | Why are some farmers profiting more from contract farming than others in Ghana? | For historical reasons, some farmer have higher self-efficacy and higher social capital | Whereas studies usually analyzed *whether* contract farming is profitable for the farmers - here the question is *for whom* it is profitable. Furthermore, the psychological concept of Self-Efficacy is introduced into agricultural economics and it is shown that it has historical roots, which makes it a cultural trait. |
| 2 | Farm Incomes | Why do some farmers in Ghana have higher incomes than others? | For historical reasons, some farmers have higher self-efficacy regarding investments, which is why they invest more and have higher incomes | Methodologically, agent based modelling and econometrics are combined. Furthermore, it is investigated how the causal effects of Self-Efficacy can be credibly identified using micro-economic and anthropological theory as well as state of the art statistical methods. The question whether cultural evolution might explain income differences amongst Ghana's pineapple farmers is also innovative. |
| 3 | Drought Adaptation | Why do not all farmers adapt to drought after they experienced it? | Farmers with higher self-efficacy adapt to drought, whereas others do so less or not at all | How well farmers adapt to adverse environmental conditions, such as droughts, is commonly explained with their socio-economic and institutional characteristics. When psychology and culture are investigated, the employed methods usually do not allow the kind of causal interpretation that is given by the authors. We demonstrate a more credible approach to test whether Self-Efficacy differences explain behavioral heterogeneity. |
| 4 | Persistent Constraints | Why is history differently persistent for different individuals? | Human and social capital, network effects, and exporting could all enable farmers to beat their historic prediction. | It is widely acknowledged that human and social capital are important for economic development and that history explains current income differences. Here, the two are brought together. We show that historically inherited constraints can be overcome with human and social capital. Thus, after many studies have established the commonality of historical persistence, we investigate how historical constraints can be relaxed. |
| 5 | Agricultural Training | Why is mulching widely diffused in Ghana and organic fertilizers are not, despite both being equally widely promoted? | Organic fertilizers are a more complex innovation than mulching. Thus, mulching can easily be learned from peers, whereas organic fertilizers require training. | Most studies find that farmers in developing countries benefit from trainings. Recently, it has been found that training are most effective to start the diffusion process but not to enhance it. We find that the effect of training depends critically on the nature of the trained technology. |
| 6 | UNESCO's World Heritage sites | Why are not all World Heritage sites promoting themselves as such? | It is especially economic and cultural incentives, whereas the collective brand equity and constraints are less important | This question could not be answered before, due to a prohibitively expensive data collection. With the development of a web-based big-data-collection-approach, over 300,000 respondents where surveyed. To efficiently use the available data, we use an innovative spatial econometric approach. |

# Chapter 1

# Explaining the Performance of Contract Farming in Ghana: The Role of Self-Efficacy and Social Capital

with Johannes Sauer[1]

March 2016

**Abstract** Self-efficacy is the belief of an individual to have the ability to be successful in a given domain. Social capital is the economic value of a person's relationships. In the context of this study, Self-efficacy is the belief of a farmer to be able to improve her income with contract farming, which increases her actual ability. Social capital increases the ability of the farmers through social support.

We surveyed 400 smallholder pineapple farmers and find that both self-efficacy and social capital are decisive for their successful integration into contract farming. To identify causal effects, we use two instruments, which are also of interest on their own: the historical presence of (1) cocoa cooperatives and (2) Christian missionary schools. During Ghana's colonial period, the British established cocoa cooperatives, which differed in their performance as a function of biogeographic factors and thus persistently shaped the self-efficacy of the farmers. Roughly at the same time, Christian missionaries established missionary schools, which impacted the traditional societies so that social capital decreased. The finding that self-efficacy and social capital are still shaped by historic variables could indicate that these variables are only slowly changing, or that they only do so in the absence of policy intervention. The latter raises the possibility that effective policies could benefit from strong reinforcing feedbacks once self-efficacy and social capital improve.

This Article is published as Wuepper, D. and J. Sauer (2016). "Explaining the Performance of Contract Farming in Ghana: The Role of Self-Efficacy and Social Capital" in Food Policy (62): 11-27.

**Contributions** David Wuepper came up with the research idea and the study design, performed the statistical analysis and wrote the article. Johannes Sauer improved the article with his feedback and suggestions throughout the whole process. Valuable input has also been contributed by Alexander Moradi, Davide Cantoni, Francesco Cinnirella, Matthias Blum, Marc Bellemare, two anonymous reviewers for Food Policy, and several conference attendees, especially at the annual conference of the International Association for Applied Econometrics.

Technical University of Munich, Agricultural Production and Resource Economics

## I. Introduction

There is an ongoing debate about the costs and benefits of contract farming for smallholders in Sub-Saharan Africa (Barrett et al. 2012, Bellemare 2012, Oya 2012, Bellemare 2015). It is a forward agreement specifying the obligations of suppliers (farmers) and buyers (processors, exporters, or supermarkets) as partners in business and widely seen as a tool for poverty mitigation, for its potential to resolve market failures (Grosh 1994). It requires the farmers to supply specified quantities and qualities and the buyers to take up the produce (often at pre-agreed prices). Additionally, the buyer commonly supplies services such as production-inputs, credit, logistics, or training (Eaton and Shepherd 2001, Will 2013).

In Ghana, contract farming has been promoted by almost all recent agricultural development projects (German Society for International Cooperation 2005, USAID 2007, 2009, Millenium Development Authority 2011, World Bank 2011, USAID 2013) for its positive, expected welfare effects (Kirsten and Sartorius 2002, Rao and Qaim 2011, Barrett et al. 2012, Bellemare 2012, Wuepper et al. 2014, Bellemare and Novak 2015). However, research has also shown important constraints to the success of contract farming (Fafchamps 1996, Fold and Gough 2008, Wuepper 2014).

As a case in point, in Ghana the performance of pineapple contract farming has been heterogeneous in time and space (Fold and Gough 2008, Barrett et al. 2012, Gatune et al. 2013) - with important socio-economic implications. The development of the pineapple export and processing sector in Ghana is directly or indirectly important for the employment and income of many. A major problem, however, is reliability. Some farmers "side-sell" fruits instead of adhering to their contracts if they can obtain a better price or faster payment locally, and farmers have reported that companies have refused to pick up fruits or pay for them when demand was unexpectedly low. These experiences had a negative effect on how farmers currently perceive contract farming.

However, some companies and farmers have apparently figured out how to make contract farming work, as indicated by the reliability and profitability of their contract agreements.

In this article, we test the hypothesis that two cultural traits, self-efficacy and social capital, explain why farmers with seemingly identical incentives and constraints are integrated into farming contracts with varying success. Both cultural traits will be discussed in the next section (II), but we will provide the following short definitions here: Self-efficacy is the belief of an individual to have the ability to achieve success in a specific domain (Bandura 1977, 1997, 2012). The concept is different from self-confidence and other related concepts and has a higher predictive and explanatory value, mainly because it is domain-specific instead of general. We define social capital following Putnam et al. (1994) as "features of social organization, such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions."

To identify the causal effects of these traits, we use "accidents of history", specifically, the colonial establishment of cocoa cooperatives and the placement of Christian missionary schools.

We find that both cultural traits are crucial for the performance of contract farming, which has important policy implications:

Self-efficacy increases how much the farmers believe to be able to benefit from contract farming, which increases their reliability, and social capital directly helps the farmers to be more reliable, e.g. by compensating for market imperfections. Policies to increase self-efficacy encourage farmers (face to face or media based) to pursue more ambitious goals (Bandura 1997, 2001, Bernard et al. 2014, 2015), support them to achieve their more ambitious goals (Bandura 1995, 1997), expose the farmers to successful peers (Bernard et al. 2014, Magnan et al. 2015), and avoid negative emotions (Bandura 2012, Haushofer and Fehr 2014, Dalton et al. 2015). Whereas these policy promise to increase the self-efficacy of the current farmer generation, it is also important to directly raise the self-efficacy of children, so that they grow up with higher levels of self-efficacy. Dercon and Singh (2013) and Dercon and Sánchez (2013) show that malnutrition during childhood persistently lowers self-efficacy in later years and Krishnan and Krutikova (2013) demonstrate in India how a specifically designed mentoring program can significantly improve the self-efficacy of poor school-children.

In the short term, important actions for policy makers and company managers is to encourage the farmers to take on more ambitious goals and to avoid failure that the farmers could attribute to their lack of ability. Furthermore, extension and trainings should not only diffuse technical knowledge but also aim at the farmers' self-efficacy – especially, it is important to avoid criticism that could make the farmers doubt their capabilities.

Policies to increase the social capital of the farmers should increase the amount of social interaction between the farmers, as demonstrated by Feigenberg et al. (2013) in India and Attanasio et al. (2009) in Colombia, and contracts must be designed to avoid trust issues, such as described by Barrett et al. (2012), so that negative experiences can be avoided.

The main contributions of our research are the identification of a cultural foundation for the performance of contract farming, an understanding of the historical roots of this cultural foundation, and a discussion of policy recommendations based on such findings.

In the next section (II), we discuss why self-efficacy and social capital matter for contract farming. In section III, we provide a succinct background of the historical sources of self-efficacy and social capital, which we later use for the identification of their effect on contract farming performance. We then turn to our data and variables in section IV and explain our empirical framework in section V. In sections VI and VII we then report our baseline and main results, respectively, and in section VIII we perform additional investigations into the effect of culture on locally generated income and participation in contract farming. We conclude our study with a discussion of our findings in section IX.

## II. Self-Efficacy and Social Capital

The performance of contract farming depends to a large extent on transaction costs. Lower transaction costs make contract farming more profitable; thus, more reliable business partners make contract farming more profitable. The following analysis is concerned with two cultural traits, one individual and one collective, that are hypothesized to affect the performance of contract farming through transaction costs. The individual trait is self-efficacy and the collective trait is social capital.

Self-efficacy is a fundamental behavior determinant that can potentially explain why some individuals are risk averse and have high discount factors in some domains. It describes how much an individual believes to have the ability to achieve success in a specific domain (Bandura 1977, 1997, 2012). It was developed originally in psychology to explain why some treatments are more helpful than others in assisting phobics with overcoming domain-specific fears (Bandura 1977). Not long after, it was discovered to explain a wide range of more common behaviors, such as educational attainments and choice of profession (Bandura 1997). Recent research in agricultural economics includes the finding that self-efficacy increases the aspirations of farmers in Ethiopia and thus motivates increased saving, credit-taking, and investments into education (Bernard et al. 2014). Whereas aspirations are only one effect of self-efficacy, it is an important one, because low aspirations caused by poverty can be a poverty trap (Moya and Carter 2014, Dalton et al. 2015). Self-efficacy can explain why poverty lowers aspirations (Bandura et al. 2001, Chiapa et al. 2012, Tafere 2014, Pasquier-Doumer and Brandon 2015) and why poverty impedes cognitive functioning, planning, and self-control (Bertrand et al. 2004, Shah et al. 2012, Mani et al. 2013, Haushofer and Fehr 2014, Laajaj 2014). Recent research in Ghana also shows that farmers with higher self-efficacy respond to adverse weather conditions with the adoption of climate-smart technology whereas others do not (Wuepper et al. 2016) and that these farmers achieve significantly higher incomes than others because they generally invest more into their fields (Wuepper and Drosten 2016).

The mechanism behind the self-efficacy effect is the following: A farmer usually only invests into a domain if she thinks it is worthwhile. Thus, it is usually insufficient for a farmer to believe that contract farming is generally profitable or has a high potential to improve welfare. Only if the farmer believes to have the ability to be increase her welfare through contract farming will she invest into it. This means her behavior is determined by what she believes to be able to achieve, not what she is objectively able to achieve (which is, however, connected, because the belief affects the outcome). Investing into contract farming can take many forms, including not side-selling when the local market price is higher; investing in quality even if quality is difficult to monitor; or adhering to the contract even if it means a lower than maximum profit in some years, all of which are for the sake of the long-term relationship with the company.

In contrast, a farmer who does not believe to have what it takes to benefit from contract farming might still enter into a contract but is likely to invest little. Even more important, once there are difficulties or temptations, it does not take much to make this farmer violate her contract.

As discussed by Bandura (1997, 2012, 2015), self-efficacy is often confused with related concepts, which are often general and not domain specific. An example is self-confidence, which is a general self-judgement. If a person has high self-efficacy in a domain that she judges to be of low value, then there might be very little positive effect from this self-efficacy on her self-confidence. Nevertheless, it might explain why the person engages in this specific domain.

It is sometimes argued that a perceived internal locus of control has an economic value (Harrison and Huntington 2000). However, imagine a person who believes to be fully in control of her life but also to lack the abilities to achieve her desired goals. This person lacks the incentive to invest effort into achieving such seemingly impossible goals (Bandura 1997).

The performance of contract farming is also likely to be affected by the social capital of the farmers. Putnam et al. (1994) define it as "features of social organization, such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions". In our context, social capital in social networks increases the reliability of farmers because it can relax constraints. Participating in a formal value chain can be riskier than selling on the local market because of investment requirements and quality standards that must be met. Social capital can increase farmers' informal access to information, labor, credit, insurance, and importantly, reduce the risk involved in contract farming, e.g. if a company does not pick up the produce or pays to late, or never. For examples of the effects of social capital see Pamuk et al. (2014), Pamuk et al. (2014), Feigenberg et al. (2013), and Guiso et al. (2008).

In the next section, we discuss the historic origins of self-efficacy and social capital.

The reason why historic events and circumstances can have very persistent effects is cultural evolution (Boyd and Richerson 1985, Henrich et al. 2008, Richerson and Boyd 2008, Boyd et al. 2011), which has often been shown to be important in economic contexts (Nunn 2009, 2012, 2013). Human decision-making is improved by our ability to imitate the behaviors of our social peers (ancestors, neighbors, family,...) instead of having to develop everything on our own. The cost of this approach is that historic circumstances that affect behavior persist to affect behavior until the behavior is re-appraised and individual learning updates the cultural knowledge. Thus, despite average efficiency gains from culture, it also implies the possibility of outdated beliefs and thus inefficient behaviors in some cases.

## III. Historical Background

Not using experimental data has advantages and disadvantages. The disadvantage is that we must worry about measurement error and unobserved heterogeneity. The advantage is that we can investigate not only the effect of cultural differences amongst the farmers but also where these cultural differences come from. These are two sides of the same coin. To exogenize the cultural traits of interest, we use "accidents of history", which are interesting on their own, as they are the long-term sources of differences in culture and contract farming performance in Ghana's pineapple sector.

Our two historical variables originate from Ghana's colonial period - then called the British Gold Coast (1878–1958). The first variable is the success rate of colonial cocoa cooperatives. After the British government abolished the transatlantic slave trade, they focused their attention on the export of agricultural commodities, such as cocoa. To improve production, they organized the cocoa farmers into cooperatives (Cazzuffi and Moradi 2010), which were in many ways similar to modern contract farming. Cooperatives were a true innovation for the approached farmers, and performance varied as much then as the performance of pineapple contract farming does today (Figure 1 in Annex A). We define the success rate of the cooperatives as the share of cooperatives in a region that survived longer than five years, which is what was recorded by the British accountants.

The second historical variable is the location of Christian missionary schools. (The location of the missionary schools can be seen in figure 2 in Annex B). Cogneau and Moradi (2011), Nunn

(2010), Woodberry (2004), and Wantchekon et al. (2015) investigate the effects of Christian missionaries and their affiliated schools. Most closely related to our research is the study by Wantchekon et al. (2015), who find that in neighboring Benin, the missionary schools persistently increased peoples' aspirations and their human capital, resulting in higher incomes today. However, in our context, we must be worried about a negative effect on social capital and possibly lower incomes. To cite two historic sources on the Gold Coast:

Ward (1966) writes about the19th-century Gold Coast

> "the introduction of Christianity and of western education brought fresh problems. Christianity and education went together, and there were inevitably many who acquired only a thin veneer. There was a good deal of trouble from semi-educated men whose scanty stock of learning led them to arrogance or downright rascality. In the early days, there was much antagonism - even sometimes rioting - between professing Christians and those who still followed the old ways,"

and Claridge (1915) reports that some missions in the Gold Coast

> "adopted a policy of separating their converts entirely from the old life for fear lest the social and artistic attractions of the old life should lead them to forget their new religion: a policy which may have been inevitable from the point of view of the Christian evangelist, but which led to a most unfortunate cleavage in the life of the community."

Both human and social capital can be important for contract farming (Eaton and Shepherd 2001; Kirsten and Sartorius 2002; Kumar and Matsusaka 2009; Barrett et al. 2012; Bellemare 2012), so the long-term effect of the missionary school placement on contract farming performance is a current gap in the literature.

## IV. Data and Variables

We representatively surveyed 400 pineapple farmers in the south of Ghana in 2013.

To be allowed to produce pineapples for export to the European Union, the farmers need to have a valid export certification which guarantees that certain production and quality standards are met (Kleemann and Abdulai 2013). Such certifications can be obtained from specialized organizations and are usually given to farm groups that are listed in the process. Thus, these lists can be used for stratified random sampling. Notably, not all certified farmers participate in contract farming, as they were often (also financially) encouraged by NGOs to participate in contract farming and many farmers did not continue on their own.

The sampling procedure was as follows: First, the major pineapple growing areas were selected and lists of groups of export-certified pineapple farmers were obtained. From these lists, farming groups were randomly selected, and several farmers were interviewed reflecting the size of their group, so that more farmers where interviewed in larger groups.

To cover non-certified farmers as well, extension agents and development agencies were asked to identify a representative sample of non-certified pineapple farmers for interviews. Such non-certified farmers might still decide to farm under contract, e.g., for Ghanaian supermarkets or fruit juice companies.

As can be seen on the maps in Annex A and B, the concentration of pineapple production close to the coast leads the sampling to be close to the coast, with the exception of Kwahu South, which is somewhat more inland.

Next, we connected our survey data to existing datasets reported previously. Murdock (1959, 1967) provides data for and locations of 834 African ethnicities as well as approximately 60 variables that describe their cultural, social, and economic characteristics. We used the data on the Ga, Akyem, Asante, Dagbami, Ewe, Fante, Grumah, and Hausa because we sampled farmers from these ethnicities. Nunn (2007) used the data from Murdock and connected it to data on the major slave trades; we use this data as well. Nunn (2010) and Cogneau and Moradi (2011) provide us with data-sets on the location of Christian missions and missionary schools; and Cazzuffi and Moradi (2010) provide us with the locations and explanatory variables on the success and failure of colonial cocoa cooperatives.

To connect Ghana's present pineapple farmers with their ancestors, we followed two strategies [see Nunn and Wantchekon (2009) for a detailed treatment]:

First, because we know the ethnicity of the sampled farmers, we can connect the farmers to their ancestors using the ethnicity information. As an example, Nunn (2007) provides data on the impact of the slave trades on the majority of ethnicities in Sub-Saharan Africa as identified by Murdock (1959). Hence, we can use the ethnicity level impact of the transatlantic slave trade as a control variable in our empirical framework, which might be important as it has been identified as a (negative) determinant of social capital in Africa, both within and across ethnicities.

Second and mostly used in this study, because we know the locations of the sampled farmers and the locations of our historical variables, we can join them together on a location basis, using GIS software:

First, we know the location of the colonial cocoa cooperatives in the 1930s, so we can count the number of successful and unsuccessful cooperatives (defined as having survived at least five years after establishment) in different radii (e.g., 5, 10, and 20 km) around our sampled farms and we can compute the success rate in the different areas which we then associate with the farmers who now reside within these areas. Secondly, we can use the locations of the missionary schools and compute how many of them were established within a 5-, 10-, and 20-km radius around today's pineapple farms.

The main variables in our following analyses are connected through locations, because they allow for more variation. Only when we control for the long-term impact of the transatlantic slave trade, do we use a variable based on the farmers' ethnicity, because Nunn and Wantchekon (2011) find that the ethnic channel is more important than the location channel, because slavery impacted culture more than institutions.

Our hypothesized channels from the historical developments are cultural too: self-efficacy and social capital, both of which are inherently difficult to capture. To capture them nevertheless, we use two different approaches:

For self-efficacy, we asked the farmers about their two main income determinants of the last two years. We then scored the answers between 1 (low self-efficacy) and 3 (high self-efficacy), depending on whether the answer included factors outside a person's control (e.g., the weather, soil, and market) or whether the answer focused on the behavior of the farmer (e.g., I learned, I improved, I adopted, I increased). Ambiguous answers were coded as a 2 (examples include yields or productivity, as these answers do not reveal how much the farmers believe to have affected the outcome). In our preferred specifications, we use the first answer of the farmers as measure for their self-efficacy, because the less time they had to think about their answer, the less likely response biases are (Bandura 1997). Our choice is informed by self-efficacy theory but there are other, plausible alternatives. Especially, it could also be argued that if a farmer states one income determinant that is under her control, she has high self-efficacy, but also using an average of both stated income determinants could be used. We tested all three variants and find a strikingly similar pattern, which suggests that farmers with high and low self-efficacy are differentiated by whether they perceive that they have a degree of influence over their income or not. Whether this is the number one factor or not does not seem to be critical.

In the analysis, we entered self-efficacy in two ways: In most specifications, we entered the variable in the form of two dummies, reflecting high and low self-efficacy. We also tested self-efficacy as a continuous variable with three values.

For social capital, we use a factor variable with three input variables: participation in social events, interpersonal trust, and number of people who would lend money to the farmer if asked. All three variables are reported by the farmers and capture different aspects of social capital. The first variable is how frequently the farmers attend social events, which include weddings, funerals, festivals and visiting church, amongst others. The attendance of social events is generally high in rural Ghana but table 1 shows that it is higher amongst contract farmers than non-contract farmers. The second variable is a generalized trust question, which asks about how much a farmers generally believes that other people can be trusted.

## Table 1. Variables and Descriptions

| VARIABLE | DESCRIPTION | CONTRACT MEAN | CONTRACT SD | NO CONTRACT MEAN | NO CONTRACT SD |
|---|---|---|---|---|---|
| INCOME | annual income from contract farming (in GHC, log in model) | 4376 | 7364 | 0 | 0 |
| PER CAPITA | annual income from contract farming per household member (in GHC, log in model) | 724 | 1188 | 0 | 0 |
| PER AREA | annual income from contract farming per acreage (in GHC, log in model) | 1106 | 2081 | 0 | 0 |
| PERCENTAGE | income share received from contract farming (in percentage) | .58 | .46 | 0 | 0 |
| COOPERATIVES | regional success rate of colonial cocoa cooperatives (% within 5 km) | .40 | .47 | .11 | .31 |
| SCHOOLS | number of Christian missionary schools around sampled farmers (w.10 km) | 17.80 | 13.53 | 15.21 | 10.18 |
| SELF-EFFICACY | open ended question on main income determinants, classified into 3 degrees of se | 2.16 | .79 | 1.78 | .76 |
| SOCIAL CAPITAL | factor variable from trust, borrow, and events | .11 | .33 | -.075 | .30 |
| TRUST | generalized trust in other people (1-6) | 2.38 | 1.87 | 2.66 | 1.79 |
| BORROW | number of people would lend the farmer money | 2.05 | 3.77 | 1.68 | 2.17 |
| EVENTS | how often the farmer attends social events in her or his village (scale 1–6) | 5.22 | 1.30 | 3.9125 | 1.86 |
| AGE | the age of the sampled pineapple farmers in 2013 | 44.01 | 10.06 | 44.52 | 11.19 |
| EDUCATION | the education level of the farmers (1–6) | 3.27 | .34 | 3.03 | .31 |
| RISK AVERSION | a farmer's preference to avoid risk; captured with a choice experiment (1–6) | 3.71 | 1.40 | 3.05 | 1.24 |
| ROADS | number of roads around a farmer's location | 4.41 | 3.85 | 3.50 | 5.08 |
| COMPANY DIST. | distance from the farms to the next company (km) | 28.36 | 30.50 | 57.86 | 34.55 |
| CITY DIST. | distance from the farms to the next city (km) | 31.86 | 7.64 | 39.80 | 19.52 |
| ACCRA DIST. | distance from the farms to the capital (km) | 37.38 | 27.46 | 63.38 | 40.63 |
| COAST DIST | distance from the farms to the coast (km) | 22.08 | 12.33 | 25.11 | 40.00 |
| TENURE SEC. | how secure the farmer believes his fields to be (1–6) | 5.58 | 1.09 | 5.48 | 1.09 |
| FARMSIZE | total land available to the farmer (in hectares) | 15.76 | 14.92 | 7.60 | 7.34 |
| TRAINING | repeated training (at least three times) (1/0) | .18 | .39 | .07 | .27 |
| PRICE PREMIUM | price differential between local and company | .09 | .068 | -.02 | .23 |
| RAINFALL | reported rainfall quantity (1–6) | 4.71 | 1.09 | 4.32 | 1.43 |
| RAIN VOLAT | squared difference between annual rainfall | 220 | 832 | 421 | 1240 |
| SOIL FERTILITY | reported fertility of the fields (1–5) | 5.26 | .81 | 5.04 | 1.06 |
| ELEVATION | elevation of the farmer's region (in m) | 82.68 | 45.16 | 86.77 | 70.43 |
| RUGGEDNESS | standard deviation of the terrain (in m) | 41.74 | 28.90 | 42.43 | 43.39 |
| SLAVERY | number of slaves exported per ethnicity (1000s) | 12309 | 1964 | 13313 | 1986 |
| RAINFALL31 | local rainfall for cocoa farms in 1931 (in mm) | 3.41 | 1.41 | 3.59 | 2.16 |
| COCOA_SOIL | soil suitability of farms for cocoa in 1931 (in %) | .51 | .49 | .16 | .36 |
| NEIGHBORS | success rate of neighboring cocoa cooperatives | .16 | .18 | .27 | .16 |
| RAILROAD | historic distance farms and railroad tracks (km) | 4376 | 7364 | .51 | 7.74 |

The question was "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people? (1= you cannot be too careful; 6= most people can be trusted)". Generalized trust is generally low in rural Ghana, which Nunn and Wantchekon (2011) argue is a long-term effect of the transatlantic slave trade.

Interestingly, trust is slightly lower amongst contract farmers than non-contract farmers, which is in line with the finding of Meijerink et al. (2014) who find evidence that participation in formal value chains can lower interpersonal trust. The final social capital variable is the number of people the farmer reports would lend her money if needed. This number is higher for contract farmers, which suggests that contract farmers perceive that others trust them more. The three variables thus capture how much the farmers interact, how much they trust others, and how much they perceive others to trust them.

We use four different measures for the performance of contract farming. The first measure is the log of income that the farmers obtain from contract farming. The second measure is the log of per capita income from contract farming, and the third measure is the per acreage income from contract farming. For the fourth measure, we use the share of income from contract farming on the total income from selling pineapples, in percent. The different income measures are dependent, as production quantities are sufficiently low in Ghana, so that companies and local markets compete for fruits. Thus, a higher income from contract farming also usually means an income higher share from contract farming. However, measuring performance as absolute income has the advantage of putting more focus on the immediately most-relevant economic indicator, whereas income share is more interesting for the intermediate to long term, as it generally captures how much company managers and farmers value the contract relationship. Currently, the sector is threatened by low pineapple supply. If the farmers only wish to sell a small share—or less—of their fruits to the companies (e.g., because this channel is perceived to be risky or no more profitable than the local market), or if the companies only wish to buy a small share—or less—(e.g., because of low quality), the profitability and sustainability of the pineapple value chain is threatened. In contrast, if farmers want to sell most—or all—of their fruits to the companies and the companies want to buy them, then we have a revealed preference for contract farming and a much better foundation for future business success. To be able to use logs when there are farmers with no income from contract farming, we added 0.01 GHC to all incomes. For the analysis, this means that we treat farmers with no contract farming income as only quantitatively different from farmers with a low contract farming income. We judge this appropriate in the context of this study, as many farmers frequently drop in and out of contracts. To see the effect of having a contract, we include a fixed effect for this variable in some of our specifications and also estimate a two part model of contract participation (discrete) and income share (continuous).

Finally, it is important that we include in some specifications the price premium that the companies offer relative to the local market prices. Controlling for the price premium could possibly lead to a downwards bias of the estimated self-efficacy effect, if it is an outcome of the farmers' self-efficacy. However, we wish to see its effect because it is plausibly affected by exogenous variables such as transportation costs and could increase the farmers' self-efficacy. To construct this variable, we calculate the average price each company offers in each location for each variety and compare this to the average price that is paid by each local market for each variety.

## V. Empirical Framework

In this section, we briefly outline our main models.

First, we regress different measures for the success of a farming contract $Y_i$ on various explanatory variables $x_i^Z$ and two cultural variables $x_i^C$, namely self-efficacy and social capital:

$$Y_i = \beta_0 + \beta_1 x_i^C + \beta_2 x_i^Z + \varepsilon_i \quad (1)$$

Then, we exogenize our cultural variables self-efficacy ($x_i^{se}$) and social capital ($x_i^{sc}$) using historical variables, namely the performance of colonial cocoa cooperatives ($x_i^{cc}$) and the placement of Christian missionary schools ($x_i^{ms}$) and estimate two stages least squares (2sls):

$$x_i^{se} = \beta_0 + \beta_1 x_i^{cc} + \beta_2 x_i^{ms} + \beta_3 x_i^z + \varepsilon_i \quad (2a)$$

$$x_i^{sc} = \beta_0 + \beta_1 x_i^{cc} + \beta_2 x_i^{ms} + \beta_3 x_i^z + \varepsilon_i \quad (2b)$$

$$Y_i = \beta_0 + \beta_1 \hat{x}_i^{se} + \beta_1 \hat{x}_i^{sc} + \beta_3 x_i^z + \varepsilon_i \quad (2c)$$

We always cluster standard errors at the district level and include district fixed effects.

Finally, we also exogenize our historical variables using a second set of instruments $x_i^I$ (based on biogeography and locations) and estimate a sequential IV model, starting with an OLS to predict exogenous variation in $x_i^{cc}$

$$x_i^{cc} = \beta_0 + \beta_1 x_i^I + \beta_2 x_i^Z + \varepsilon_i \tag{3}$$

The placement of Christian missionary schools can be taken as exogenous but we need the instruments $x_i^I$ to exogenize the performance of the colonial cocoa cooperatives. We use the rainfall on the cocoa farms, their soil suitability for cocoa and their distance to the colonial railroad tracks for $x_i^I$. As we argue below, all three variables affected the reliability of the cocoa farmers but did not have any other long-term effect on the current generation of pineapple farmers.

We then estimate another 2SLS in which we use the predicted $\hat{x}_i^{cc}$ from (3) as instrument:

$$x_i^{se} = \beta_0 + \beta_1 \hat{x}_i^{cc} + \beta_2 x_i^{ms} + \beta_3 x_i^z + \varepsilon_i \tag{4a}$$

$$x_i^{sc} = \beta_0 + \beta_1 \hat{x}_i^{cc} + \beta_2 x_i^{ms} + \beta_3 x_i^z + \varepsilon_i \tag{4b}$$

$$Y_i = \beta_0 + \beta_1 \hat{x}_i^{se} + \beta_1 \hat{x}_i^{sc} + \beta_3 x_i^z + \varepsilon_i \tag{4c}$$

To take into account the measurement error from the "first stage" we bootstrap the standard errors in (4) but also report regular standard errors as comparison. In addition to several specifications of the main models in which the amount of control variables is systematically varied, we also estimate a range of supporting models to investigate our main identifying assumptions.

## VI. Baseline Results

We start our analysis by testing for an empirical relationship between self-efficacy, social capital and pineapple contract farming performance.

We always estimate several specifications, starting without any control variables and then increasing their number. The basic idea is to learn about the relative risk of omitted variable bias on the one hand, and the inclusion of endogenous control variables on the other.

For brevity, we only show the first and the last specification – the first being more at risk of omitted variable bias and the latter being more at risk of having endogenous control variables.

**Table 2. Testing the Relationship between Culture and Contract Farming (A)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | OLS | OLS | OLS | OLS |
| DEP.VAR | INCOME | INCOME | PERCENTAGE | PERCENTAGE |
| HIGH SELF-EFFICACY | 1.633** | 1.680*** | 0.143** | 0.138*** |
| | (0.524) | (0.296) | (0.0438) | (0.0292) |
| LOW SELF-EFFICACY | -0.726 | -0.243 | -0.0602 | -0.0259 |
| | (0.550) | (0.439) | (0.0441) | (0.0331) |
| SOCIAL CAPITAL | 0.728*** | 0.362*** | 0.0657*** | 0.0439*** |
| | (0.146) | (0.102) | (0.00793) | (0.00625) |
| CONTROLS | NO | YES | NO | YES |
| CONTRACT | NO | YES | NO | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.31 | 0.46 | 0.38 | 0.54 |
| N | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls include age and education, household size, certification, price premium paid by the companies, farm size, distance to the capital, rainfall amount and volatility, soil quality, distance to coast, terrain ruggedness, and roads. Contract controls for whether a farmer has a formal contract with a company or not.

Controls include age and education of the farmer, household size, whether she is certified, whether the has a contract arrangement, the price premium offered by the local companies, her land that she uses for the production of pineapples, the distance to the capital, cities, and the coast, rainfall and rainfall volatility, the ruggedness of the terrain, and the local road infrastructure. In all specifications, we include district fixed effects. Standard errors are clustered at the district level.

Table 2 indicates a robust and significant, positive empirical relationship between high-self-efficacy and contract farming performance and a negative, but insignificant, empirical relationship between low self-efficacy and contract farming performance. This is independent of whether contract farming performance is defined as log of income or percentage share of income. Similarly, the indicated relationship between social capital and contract farming performance is robust, significant, and positive.

In table 3, we report the same results but for differently defined contract farming variables. Whereas contract farming performance was operationalized in table 2 as log of income from contract farming (spec. 1 and 2) and as percentage income from contract farming (spec. 3 and 4), in table 3, it is log of income per capita (spec. 1 and 2) and per acreage (spec. 3 and 4). The pattern is strikingly similar across all specifications.

**Table 3. Testing the Relationship between Culture and Contract Farming (B)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | OLS | OLS | OLS | OLS |
| DEP.VAR | PER CAPITA | PER CAPITA | PER AREA | PER AREA |
| HIGH SELF-EFFICACY | 1.091** | 1.124*** | 1.148** | 1.169*** |
| | (0.350) | (0.197) | (0.371) | (0.219) |
| LOW SELF-EFFICACY | -0.501 | -0.164 | -0.543 | -0.200 |
| | (0.381) | (0.303) | (0.421) | (0.322) |
| SOCIAL CAPITAL | 0.469*** | 0.209** | 0.525*** | 0.290*** |
| | (0.0984) | (0.0715) | (0.0776) | (0.0694) |
| CONTROLS | NO | YES | NO | YES |
| CONTRACT | NO | YES | NO | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.19 | 0.45 | 0.29 | 0.45 |
| N | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls include age and education, household size, certification, price premium paid by the companies, farm size, distance to the capital, rainfall amount and volatility, soil quality, distance to coast, terrain ruggedness, and roads. Contract controls for whether a farmer has a formal contract with a company or not.

## VII. Main Results

There are multiple reasons why we cannot interpret the baseline results as causal effects. First of all, self-efficacy is a mental model. People habitualize what they try to achieve, how much they try to achieve, and how persistently they try to achieve it. This can be self-limiting if it leads to low levels of effort in low-level pursuits, when more would have been possible. The danger is that we capture not only self-efficacy but also other incentives and constraints that affect our measure of self-efficacy and the performance of contract farming. Examples are farmers who live in areas with better rainfall or who own fields closer to a company. They might have a high degree of self-efficacy because the challenge they face is smaller than that of other farmers. In the extreme, our self-efficacy measure exclusively captures unobserved incentives and constraints and has no additional explanatory value once a full model is specified.

Similarly, social capital might not only cause contract farming success, but perhaps a profitable business can also bring together farmers and foster the development of social capital.

To exogenize our cultural variables, we use "accidents of history" as sources of exogenous variation. This has the additional advantage that we might learn more about the reason why we observe different levels of self-efficacy and social capital in the first place. This could sometimes have consequences for the recommended policies.

We use the success rate of colonial cocoa cooperatives as an instrument for self-efficacy in the domain of a formal value chain. These cooperatives were established basically for the same reasons as why development agencies currently promote contract farming: intensification of production, leading to better quality and larger quantities. By organizing and bundling production, the colonial government wanted to make it easier to provide inputs and services to the farmers and to monitor quality, which was crucial for the export market (Cazzuffi and Moradi 2010). According to Cultural Evolution Theory (Boyd and Richerson 1985, Richerson and Boyd 2008, Boyd et al. 2011), the historic experiences with the cocoa cooperatives could have shaped the beliefs of the farmers and those of their children. The performance of the colonial cocoa cooperatives is unlikely to have a direct effect on the performance of current pineapple contract farming. One reason is that no farmer in our sample is involved in cocoa production. The farmers in our sample grow cassava, corn, yams, cowpea, and other crops – mainly for personal consumption – and pineapples – mainly as their sole cash crop. A potential threat is any economic advantage that would have accrued from successful cocoa cooperatives and that would still make some communities economically better off than others. This would e.g., raise the local market price (making contract farming less attractive), or it could improve factor markets and the ability of the farmers to comply with contract requirements (making contract farming easier). This hypothesis can be tested. In table 4, we show two specifications which attempt to explain the local market price with the performance of the colonial cocoa cooperatives. The relationship is always insignificant.

**Table 4. The long-Term Wealth Effect of the Cocoa Coooperatives**

| SPEC. | (1) | (2) |
|---|---|---|
| MODEL | OLS | OLS |
| DEP.VAR | LOCAL PRICE | LOCAL PRICE |
| COOPERATIVES | 0.00263<br>(0.00266) | -0.00159<br>(0.00165) |
| CONTROLS | RAINFALL,<br>TOPOGRAPHY, | DISTANCES TO<br>COMPANIES AND CITIES |
| DISTRICT FE | YES | YES |
| R-SQ | 0.278 | 0.273 |
| N | 398 | 398 |

Notes: Standard errors are clustered at the district level.
Significance levels are 10% (*), 5% (**), and 1% (***).

Another threat comes from a possible long-term persistence of self-efficacy. Since we argue that self-efficacy is important for contract farming, it logically follows that it might have been important for the colonial cooperatives. Possibly, both the performance of colonial cocoa cooperatives and of current pineapple contracts is determined by persistent differences in self-efficacy. Thus, we must later also find instruments to exogenize the performance of the cooperatives in order to identify their causal effect on the farmers' self-efficacy.

To instrument the social capital of Ghana's pineapple farmers, we use the placement of Christian missionary schools, which might have interrupted village social relationships and traditional networks, as discussed above.

As Cogneau and Moradi (2011) and Nunn (2010) describe, the location of the missions themselves were influenced by several factors including disease, environment, and existing infrastructure. Acemoglu et al. (2014) however, argue that Christian missions are valid instruments, amongst other reasons, because different missions had different strategies, which balance each other out on average. This pattern is also obvious in Ghana, where some missionaries went to especially poor areas in order to have the largest impact, whereas others went to especially safe and productive areas to make their lives a little easier.

For our research, we are using Christian missionary schools which do not require a balancing out of different strategies because their locations were determined quite idiosyncratically (it was typically an individual teacher who could decide where to start a school and many of the local teachers simply went back to their own communities) (see also: Macdonald (1898), Claridge (1915); Ward (1966); Nunn (2010); Wantchekon et al. (2015) for similar evaluations).

**Table 5. Explaining the Locations of Christian Missionary Schools**

| SPEC. | (1) | (2) | (3) |
|---|---|---|---|
| MODEL | OLS | OLS | OLS |
| DEP.VAR | SCHOOLS | SCHOOLS | SCHOOLS |
| RAINFALL | 0.0176 | -0.00532 | 0.0184 |
| | (0.0144) | (0.0120) | (0.0169) |
| SOIL PROBLEMS | -0.0494 | -0.0213 | -0.0601 |
| | (0.0583) | (0.0129) | (0.0413) |
| COAST DISTANCE | -0.0794 | -0.0778 | -0.212 |
| | (0.568) | (0.324) | (0.309) |
| RUGGEDNESS | 0.504*** | 0.634*** | 0.516*** |
| | (0.138) | (0.0738) | (0.0627) |
| SLAVERY | | -0.0662 | |
| | | (0.0414) | |
| SUBSISTENCE TREES | | -0.0557 | |
| | | (0.139) | |
| SUBSISTENCE CEREALS | | -0.0595 | |
| | | (0.144) | |
| MISSIONS | | 0.947*** | |
| | | (0.251) | |
| SLAVERY UNHEREDITARY | | | -0.374 |
| | | | (0.290) |
| SLAVERY HEREDITARY | | | -0.0755 |
| | | | (0.0832) |
| RIVERS | | | 0.691 |
| | | | (0.473) |
| AREA DEVELOPMENT | | | -0.240 |
| | | | (0.280) |
| DISTRICT FE | YES | YES | YES |
| R-SQ | 0.65 | 0.89 | 0.69 |
| N | 398 | 385 | 398 |

Notes: Standard errors are clustered at the district level.
Significance levels are 10% (*), 5% (**), and 1% (***).

To probe the randomness of Christian missionary school placement, table 5 presents three basic regressions that attempt to explain the number of schools in a region. It can be seen that rainfall and poor soils neither attracted nor repelled schools, and similarly, distance to the coast, impact of the transatlantic slave trade, the main subsistence crop, the prevailing form of slavery, the amount of rivers, and regional development (a combination of settlement structure and political centralization, created from the Murdock data) cannot explain why schools were established where they were. Only the ruggedness of the terrain (the standard deviation of the elevation in a region) and the number of missions in the region are significant explanations for school placement. This suggests that Christian missionary schools can be treated as exogenously given.

Nevertheless, we must worry about the schools' long term effect on educational attainments, which could influence contract farming performance directly. For this reason, we always control for the education of the farmers. However, we do not find that educational differences generally explain differences in contract farming performance, so we conclude that Christian missionary schools do not cause any selection bias in our sample.

Before we have a closer look at the determinants of colonial cooperative performance, we start our main analysis by testing for an empirical relationship between the success rate of the colonial cocoa cooperatives (cooperatives) and the placement of Christian missionary schools on the one hand and the current performance of pineapple contract farming on the other.

If we cannot find that cooperatives and schools are robustly correlated with current contract farming performance, then there is no value in working with them as instruments. However, as table 6 and 7 show, the empirical relationships are strong and do not look too different from the relationships that are presented in tables 2 and 3 between self-efficacy and social capital on the one hand and pineapple contract farming performance on the other.

Table 6 indicates a strong persistence between the success of the colonial cocoa cooperatives and current pineapple contract farming, irrespective of whether the latter is defined as log of income (specifications 1 and 2) or as percentage income share from contract farming (specifications 3 and 4) or whether few or many covariates are included in the specification.

An equally robust - but negative - empirical relationship can be found between contract farming performance and the historic locations of Christian missionary schools.

**Table 6. Testing the Relationship between History and Contract Farming (A)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | OLS | OLS | OLS | OLS |
| DEP.VAR | INCOME | INCOME | PERCENTAGE | PERCENTAGE |
| COOPERATIVES | 1.801*** | 1.641*** | 0.239*** | 0.214*** |
| | (0.237) | (0.234) | (0.0464) | (0.0447) |
| SCHOOLS | -1.031*** | -0.964*** | -0.145*** | -0.149*** |
| | (0.293) | (0.289) | (0.0425) | (0.0344) |
| CONTROLS | NO | YES | NO | YES |
| CONTRACT | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.24 | 0.39 | 0.29 | 0.45 |
| N | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).. Controls include age and education, household size, certification, price premium paid by the companies, farm size, distance to the capital, rainfall amount and volatility, soil quality, distance to coast, terrain ruggedness, roads, and the impact of the trans-Atlantic slave trade. Contract is a fixed effect for having a contract arrangement.

Table 7 also shows this pattern, with the difference that the dependent variable is now either defined as log of contract farming income per capita (specification 1 and 2) or per hectare (specifications 3 and 4).

We always estimate several specifications, starting without any control variables and then increasing their number. In table 6 and 7, we only show the first and the last specifications. The latter includes controls for age and education of the farmer, her household size, whether she is certified, the price premium offered by the local companies, her land that she uses for the production of pineapples, the distance to the capital and to the coast, rainfall and rainfall volatility, the ruggedness of the terrain, the local road infrastructure, and as a control for another, potentially important historical event, the impact of the transatlantic slave trade.

To summarize what we have found so far, the performance of current pineapple contracts is positively correlated with self-efficacy, social capital, and the performance of colonial cocoa cooperatives and negatively correlated with the historic locations of Christian missionary schools.

**Table 7. Testing the Relationship between History and Contract Farming (B)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | OLS | OLS | OLS | OLS |
| DEP.VAR | PER CAPITA | PER CAPITA | PER AREA | PER AREA |
| COOPERATIVES | 1.801*** | 1.641*** | 2.026*** | 1.797*** |
| | (0.237) | (0.234) | (0.259) | (0.235) |
| SCHOOLS | -1.031*** | -0.964*** | -1.227*** | -1.094*** |
| | (0.293) | (0.289) | (0.332) | (0.282) |
| CONTROLS | NO | YES | NO | YES |
| CONTRACT. | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.23 | 0.39 | 0.25 | 0.38 |
| N | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).. Controls include age and education, household size, certification, price premium paid by the companies, farm size, distance to the capital, rainfall amount and volatility, soil quality, distance to coast, terrain ruggedness, roads, and the impact of the trans-Atlantic slave trade. Contract is a fixed effect for having a contract arrangement.

The next step is to have a closer look at the two cultural traits.

Table 8a shows the results of four regressions of self-efficacy on explanatory variables and two regressions of social capital on explanatory variables. As we discussed above, we do not need to worry about selection bias with the historic location of the Christian missionary schools, but we do need to worry about selection bias when it comes to the performance of the colonial cocoa cooperatives. Thus, in specifications 3 and 4 the performance of the colonial cocoa cooperatives is instrumented with four different instruments that are shown in detail in table 8b.

We always start with clearly exogenous explanatory variables and then add increasingly endogenous variables. In table 8a, we show two specifications for each model, one with fewer explanatory variables and one with more. It can be seen that self-efficacy is not affected by rainfall, infrastructure, or price level, which are all factors that objectively increase the economic ability of a farmer. Despite historical circumstances, tenure security is the only current context variable that has a significant relationship with self-efficacy. The strongest and most robust explanation for self-efficacy is the success of colonial cocoa cooperatives. Once this success is exogenized, using instrumental variables in specification 3 and 4, the causal effect is stronger than the observed correlation in specifications 1 and 2.

**Table 8a. Understanding The Cultural Traits**

| SPEC. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| MODEL | OLS | OLS | 2SLS | 2SLS | OLS | OLS |
| DEP.VAR | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SOCIAL CAPITAL | SOCIAL CAPITAL |
| COOPERATIVES | 0.138** | 0.166*** | 0.180*** | 0.175*** | 0.00600 | -0.00521 |
| | (0.0440) | (0.0206) | (0.0336) | (0.0479) | (0.0251) | (0.0610) |
| SCHOOLS | -0.0401 | -0.0925** | -0.101** | -0.0600 | -0.275*** | -0.356*** |
| | (0.0721) | (0.0345) | (0.0459) | (0.0575) | (0.0486) | (0.0532) |
| ROADS | -0.000295 | 0.0861 | 0.0883 | 0.00609 | | |
| | (0.0306) | (0.0979) | (0.0936) | (0.0393) | | |
| RAIN VOLATILITY | -0.0672* | -0.0686* | -0.0619** | -0.0748** | -0.0330 | -0.0199 |
| | (0.0322) | (0.0328) | (0.0305) | (0.0296) | (0.0573) | (0.0541) |
| PRICE | | 0.0420 | | 0.0334 | | |
| | | (0.0352) | | (0.0432) | | |
| COMPANY DISTANCE | | 0.152 | 0.147 | | | |
| | | (0.189) | (0.171) | | | |
| TENURE SECURITY | | 0.0682** | | 0.0654*** | | |
| | | (0.0268) | | (0.0239) | | |
| MANY TRAININGS | | | | | | -0.434* |
| | | | | | | (0.221) |
| SLAVERY | | | | | | -0.134** |
| | | | | | | (0.0513) |
| AGE AND EDUCATION | YES | YES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES | YES | YES |
| R-SQ | 0.18 | 0.20 | 0.18 | 0.20 | 0.11 | 0.17 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).

The strongest and most robust explanation for social capital is the historic placement of Christian missionary schools. We also find empirical relationships with the reception of multiple trainings from development organizations (positive) and the impact of the transatlantic slave trade (negative). However, both variables are plausibly endogenous, so this is at best indicative.

Turning to the first stages that make the performance of the colonial cocoa cooperatives exogenous in specifications 3 and 4, we must discuss why our instruments work.

Cazzuffi and Moradi (2010) analyze the performance of the colonial cocoa cooperatives and provide explanatory variables that include feasible instruments. First, there is the local rainfall data for the year 1931. Conditional on controlling for local rainfall in 2013, it is uncorrelated with the current pineapple contract farming performance but does explain variation in the cocoa cooperative performance.

**Table 8b. First Stages for Specifications (3) and (4) of Table 8a**

| SPEC. | (3) | (4) |
|---|---|---|
| MODEL | 2SLS | 2SLS |
| DEP.VAR | COOPERATIVES | COOPERATIVES |
| RAINFALL 1930S | 1.011*** | .831*** |
| | (.152) | (.159) |
| COCOA SOIL | .528*** | .692*** |
| | (.158) | (.219) |
| COCOA NEIGHBORS | .078*** | .224*** |
| | (.026) | (.033) |
| DISTANCE RAILROAD | -.322*** | -.659* |
| | (.144) | (.369) |
| ALL COVARIATES OF 2ND STAGE | YES | YES |
| DISTRICT FE | YES | YES |
| F EXCLUDED | 57.66 | 83.96 |
| R-SQ | 0.87 | 0.92 |
| N | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).

A concern could be climate change, which could have changed the spatial distribution between 1931 and 2013. However, in large parts of our sampling region, climate change did not substantially change the spatial rainfall distribution.

Our second instrument is the cocoa soil suitability data for the year 1931 which, controlling for pineapple soil suitability in 2013, works as an instrument similar to the rainfall variable but perhaps even more robustly so. This is the instrument for which the exclusion restriction holds almost certainly because cocoa and pineapple clearly have different soil requirements.

Third, we also use the success rate of the colonial cocoa cooperatives of neighboring villages, which we identify in geographic information software (GIS). This removes persistent, local influences (assuming they were not present in the neighboring villages too) and is based on the assumption that those social variables that operated at the village level and that affected the success-rate of the cooperatives can be removed by using the values from adjacent villages (Bramoullé et al. 2009).

Finally, we also use GIS to identify the distance between the farms and the historic railroad. Cocoa was transported by train to the coast whereas today, pineapples are transported by truck to the companies north of Accra. Thus, we argue that railroad distance explains why transaction costs were exogenously varied for the cocoa farmers, which created exogenous variation in their performance, but it has no other effect on the current performance of the pineapple contract. The Achilles' heel of this instrument is possible long-term income effects from distance to the railroad (Jedwab and Moradi 2012, 2015). As a test, we investigated, as before, whether historic railroad distance explains local market prices today and find no empirical relationship. This suggests that pineapple farmers are only affected by distance to the railroad in as much of the fact that such distance impacted the success probability of the cocoa cooperatives during colonial times.

Table 8b shows that all of our instruments are strongly correlated with the performance of the colonial cocoa cooperatives, and the F test indicates that we do not need to worry about weak instrument bias.

In conclusion, we find that the performance of the colonial cocoa cooperatives is the best available explanation for the degree of self-efficacy of the farmers and the historic location of Christian missionary schools is the best available explanation for the social capital in the villages. Self-efficacy is important to motivate farmers to sufficiently invest in the success of contracts and to keep on investing in the face of adversity instead of giving up or being easily tempted into violating the contract. Community social capital on the other hand increases the ability of the farmer to be a reliable business partner because it works as a social safety net if something goes wrong and can be a source of information, labor, credit, and other important inputs.

In the next step, we use our historical variables to instrument self-efficacy and social capital in order to identify their causal effect on the performance of contract farming. We start with the first stages in tables 9a (self-efficacy) and 9b (social capital) and show the results of the second stages in table 9c (absolute income) and 9d (income share). Afterwards, we address the issue of migration between colonial times and the present in table 10, by excluding regions with higher migration rates. In table 11 we show the results of instrumenting also the performance of the colonial cooperatives and in tables 12a to 12c we show the results of a following 2SLS.

**Table 9a. 2SLS First Stage – The Historic Determinants of Self-Efficacy**

| SPEC. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY |
| COOPERATIVES | 0.476*** (0.0924) | 0.584*** (0.0967) | 0.564*** (0.0941) | 0.461*** (0.0723) | 0.588*** (0.0945) | 0.568*** (0.0942) |
| SCHOOLS | -0.246*** (0.0591) | -0.406*** (0.0978) | -0.408*** (0.0968) | -0.256*** (0.0707) | -0.386*** (0.105) | -0.387*** (0.105) |
| CONTRACT | YES | YES | YES | NO | NO | NO |
| COVARIATES OF 2ND STAGE | YES | YES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES | YES | YES |
| R-SQ | 0.23 | 0.28 | 0.28 | 0.22 | 0.28 | 0.28 |
| F EXCLUDED | 13.32 | 22.87 | 21.21 | 25.55 | 36.82 | 19.07 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). "Contract" is a fixed effect for having a contract arrangement.

As can be seen in tables 9a and 9b, the cocoa cooperatives and Christian missionary schools are strong instruments for farmers' self-efficacy and social capital, as judged by the F test shown at the bottom of the tables. However, we need to justify why we think that the exclusion restrictions hold. In the beginning of this section, we briefly discussed several concerns and argued that they are unlikely to create selection bias. Subsequently, we tested whether the performance of the colonial cocoa cooperatives and the current performance of the pineapple contracts might be both caused by persistent degrees of self-efficacy (table 8) and found that the difference in overall self-efficacy and instrumented self-efficacy is economically not significant.

**Table 9b. 2SLS First Stage – The Historic Determinants of Social Capital**

| SPEC. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL |
| COOPERATIVES | 0.213*** | 0.245*** | 0.275*** | 0.161*** | 0.235*** | 0.268*** |
| | (0.0429) | (0.0328) | (0.0676) | (0.0377) | (0.0421) | (0.0627) |
| SCHOOLS | -0.530*** | -0.672*** | -0.676*** | -0.564*** | -0.717*** | -0.716*** |
| | (0.0308) | (0.0629) | (0.0588) | (0.0489) | (0.0436) | (0.0439) |
| COVARIATES OF 2ND STAGE | YES | YES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES | YES | YES |
| R-SQ | 0.20 | 0.24 | 0.25 | 0.16 | 0.24 | 0.24 |
| F EXCLUDED | 215.65 | 43.28 | 41.89 | 78.13 | 142.38 | 173.13 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). “Contract” is a fixed effect for having a contract arrangement.

Thus, we find that cooperatives and missionary schools are feasible instruments to identify the true causal effect of self-efficacy and social capital on contract farming performance in table 9c. Tables 9c and 9d present the empirical evidence that self-efficacy and social capital have significant and causal effects on the income obtained from pineapple contract farming (table 9c) and on the percentage share of income obtained from pineapple contract farming compared to income obtained from the local market (table 9d).

**Table 9c.2SLS 2nd Stage - The Determinants of Contract Farming Income**

| SPEC. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | CONTRACT INCOME | CONTRACT INCOME | CONTRACT INCOME | CONTRACT INCOME | CONTRACT INCOME | CONTRACT INCOME |
| SELF-EFFICACY | 4.997*** (0.500) | 5.005*** (0.476) | 3.899*** (0.609) | 4.562*** (0.517) | 4.902*** (0.535) | 3.842*** (0.662) |
| SOCIAL CAPITAL | 1.565*** (0.324) | 1.732*** (0.530) | 2.427*** (0.505) | 1.929*** (0.227) | 1.910*** (0.384) | 2.519*** (0.451) |
| RISK PREF | | | 0.481** (0.243) | | | 0.717* (0.389) |
| NETWORK | | | 0.482*** (0.186) | | | 0.795*** (0.305) |
| RAINFALL | 1.227*** (0.463) | 1.081** (0.491) | 0.923** (0.370) | 1.135*** (0.378) | 1.049** (0.456) | 0.907** (0.361) |
| RAIN VOLATILITY | 0.635*** (0.237) | 0.771*** (0.280) | 0.516* (0.302) | 0.582** (0.253) | 0.770*** (0.287) | 0.514* (0.308) |
| ELEVATION | 0.680 (0.283) | 0.462 (0.256) | 0.136 (0.492) | 0.538 (0.455) | 0.412 (0.414) | 0.110 (0.501) |
| RUGGEDNESS | -1.323*** (0.388) | -1.411** (0.550) | -1.182* (0.639) | -1.342*** (0.362) | -1.374** (0.571) | -1.163* (0.645) |
| COMPANY DISTANCE | | 0.269 (0.474) | 0.461 (0.537) | | 0.317 (0.470) | 0.486 (0.541) |
| CITY DISTANCE | | -0.142 (0.534) | -0.191 (0.490) | | -0.179 (0.519) | -0.209 (0.475) |
| PRICE PREMIUM | | -0.407 (0.676) | -0.207 (0.595) | | -0.344 (0.700) | -0.175 (0.597) |
| CERTIFIED | | 0.828** (0.387) | 0.523 (0.418) | | 0.890** (0.358) | 0.550 (0.399) |
| LAND | | 0.673*** (0.215) | 0.664*** (0.233) | | 0.678*** (0.210) | 0.668*** (0.229) |
| CONTRACT | YES | YES | YES | NO | NO | NO |
| AGE, EDUCATION, HOUSEHOLD | YES | YES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | | YES | YES |
| R-SQ | 0.06 | 0.07 | 0.20 | 0.09 | 0.08 | 0.20 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). "Contract" is a fixed effect for having a contract arrangement.

**Table 9d. 2nd Stage - The Determinants of the Contract Farming Income Share**

| SPEC. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | PERCENT. | PERCENT. | PERCENT. | PERCENT. | PERCENT. | PERCENT. |
| SELF-EFFICACY | 0.422***<br>(0.0326) | 0.436***<br>(0.0317) | 0.337***<br>(0.0386) | 0.403***<br>(0.0348) | 0.425***<br>(0.0346) | 0.329***<br>(0.0407) |
| SOCIAL CAPITAL | 0.161***<br>(0.0352) | 0.128***<br>(0.0462) | 0.191***<br>(0.0431) | 0.176***<br>(0.0214) | 0.147***<br>(0.0326) | 0.202***<br>(0.0365) |
| RISK PREF | | | 0.786***<br>(0.298) | | | 0.0656**<br>(0.0333) |
| NETWORK | | | 0.716*<br>(0.399) | | | 0.0715***<br>(0.0243) |
| RAINFALL | 0.103***<br>(0.0397) | 0.0971**<br>(0.0416) | 0.0828**<br>(0.0338) | 0.0992***<br>(0.0322) | 0.0936**<br>(0.0370) | 0.0807***<br>(0.0309) |
| RAIN VOLATILITY | 0.0396<br>(0.0275) | 0.0452<br>(0.0290) | 0.0221<br>(0.0321) | 0.0374<br>(0.0285) | 0.0451<br>(0.0296) | 0.0219<br>(0.0325) |
| ELEVATION | 0.0442<br>(0.0334) | 0.0319<br>(0.0308) | 0.00230<br>(0.0331) | 0.0381<br>(0.0349) | 0.0265<br>(0.0322) | -0.00102<br>(0.0346) |
| RUGGEDNESS | -0.121***<br>(0.0293) | -0.149***<br>(0.0402) | -0.128***<br>(0.0428) | -0.122***<br>(0.0274) | -0.145***<br>(0.0420) | -0.126***<br>(0.0444) |
| COMPANY DISTANCE | | 0.0231<br>(0.0310) | 0.0403<br>(0.0370) | | 0.0282<br>(0.0309) | 0.0435<br>(0.0371) |
| CITY DISTANCE | | 0.0521<br>(0.0330) | 0.0476<br>(0.0296) | | 0.0481<br>(0.0333) | 0.0453<br>(0.0304) |
| PRICE PREMIUM | | -0.0267<br>(0.0589) | -0.00854<br>(0.0518) | | -0.0199<br>(0.0580) | -0.00453<br>(0.0498) |
| CERTIFIED | | 0.0415<br>(0.0305) | 0.0141<br>(0.0323) | | 0.0482<br>(0.0304) | 0.0175<br>(0.0329) |
| LAND | | 0.0309<br>(0.0189) | 0.0299<br>(0.0189) | | 0.0315*<br>(0.0186) | 0.0304<br>(0.0188) |
| CONTRACT | YES | YES | YES | NO | NO | NO |
| AGE, EDUCATION, HOUSEHOLD | YES | YES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES | YES | YES |
| R-SQ | 0.04 | 0.06 | 0.25 | 0.07 | 0.07 | 0.25 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). "Contract" is a fixed effect for having a contract arrangement.

A concern we have not addressed so far is migration. Since we spatially connect most of our historic variables with our cultural variables, we must worry about the effect of farmers having changed their location between colonial times and the present. To investigate this, we start by examining the data of Kleemann and Abdulai (2013), who have a 50% overlap with our sample and who included a question on migration in their survey. They find that about 25% of their sampled farmers did migrate - but usually not far. This is a typical pattern in southern Ghana, where migration is relatively common within regions or between rural and urban locations. To investigate whether migration creates a problem for our identification strategy, we re-estimate table 9 but exclude all communities in which less than 70% of the farmers have been born in the same community in which they currently live. Table 10a and 10b show the first stages. A caveat of this data is that it only covers recent migration, and not historic migration. However, recent migration is a proxy for historic migration, and in most areas, migration is higher more recently than it was historically.

We find that despite the lower sample size, the estimates look reassuringly similar to our estimates from tables 9a and 9b.

In table 10c, we present the results of the second stage. It can be seen that the estimated causal effects are slightly stronger than in table 9c, which suggests that migration possibly introduces a slight measurement error but does not otherwise bias our causal estimates.

**Table 10a. 2SLS First Stage – The Effect of Migration on the Previous Results (A)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY | SELF-EFFICACY |
| COOPERATIVES | 0.428*** | 0.563*** | 0.428*** | 0.563*** |
| | (0.0677) | (0.0937) | (0.0677) | (0.0937) |
| SCHOOLS | -0.221** | -0.351** | -0.221** | -0.351** |
| | (0.0754) | (0.110) | (0.0754) | (0.110) |
| COVARIATES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.22 | 0.29 | 0.22 | 0.29 |
| F EXCLUDED | 29.11 | 33.54 | 29.11 | 33.54 |
| N | 277 | 277 | 277 | 277 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).

**Table 10b. 2SLS First Stage – The Effect of Migration on the Previous Results (B)**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL | SOCIAL CAPITAL |
| COOPERATIVES | 0.113*** (0.0231) | 0.167*** (0.0375) | 0.113*** (0.0231) | 0.167*** (0.0375) |
| SCHOOLS | -0.530*** (0.0539) | -0.669*** (0.0476) | -0.530*** (0.0539) | -0.669*** (0.0476) |
| COVARIATES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.19 | 0.26 | 0.19 | 0.26 |
| F EXCLUDED | 56.23 | 109.95 | 56.23 | 109.95 |
| N | 277 | 277 | 277 | 277 |

Notes: Standard errors are clustered at district level. Significance levels are 10% (*), 5% (**), and 1% (***).

**Table 10c. 2SLS Second Stage – The Effect of Migration on the Previous Results**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | INCOME | INCOME | PERCENTAGE | PERCENTAGE |
| SELF-EFFICACY | 5.050*** (0.557) | 4.217*** (0.486) | 0.449*** (0.0246) | 0.448*** (0.0371) |
| SOCIAL CAPITAL | 2.042*** (0.259) | 1.751*** (0.487) | 0.185*** (0.0196) | 0.148*** (0.0290) |
| COVARIATES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.16 | 0.20 | 0.156 | 0.216 |
| N | 277 | 277 | 277 | 277 |

Notes: Standard errors are clustered at district level. Significance levels are 10% (*), 5% (**), and 1% (***).

Finally, we need to test whether the performance of the colonial cocoa cooperatives is actually exogenous. The reason is the possibility of a reverse causality. We argue that the performance of colonial cocoa cooperatives shaped the subsequent self-efficacy of the farmers. An alternative could be that already the performance of the colonial cocoa cooperatives were determined by differences in self-efficacy and these differences persist until today. This of course would suggest that self-efficacy is much more stable and thus less responsive to policy. In table 8, we already presented evidence that the colonial cocoa cooperatives performance was a function of several instruments, including biogeography and historic locations. We briefly summarize three instruments and then use them in an OLS regression (table 11), to predict exogenous variation in the performance of the colonial cocoa cooperatives, which we then use in another 2SLS regression (tables 12a,b, and c).

**Table 11. Exogenizing the Success of the Colonial Cocoa Cooperatives**

| SPEC. | (1) | (2) | (3) |
|---|---|---|---|
| MODEL | OLS | OLS | OLS |
| DEP.VAR. | COOPERATIVES | COOPERATIVES | COOPERATIVES |
| RAINFALL 1930S | 0.995*** | 0.816*** | 0.803*** |
| | (0.0788) | (0.154) | (0.157) |
| COCOA SOIL | 0.411*** | 0.410*** | 0.411*** |
| | (0.0621) | (0.0631) | (0.0576) |
| DISTANCE RAILROAD | -0.00198 | -0.343** | -0.339** |
| | (0.0284) | (0.143) | (0.141) |
| ALL COVARIATES OF 2ND STAGE | YES | YES | YES |
| DISTRICT FE | YES | YES | YES |
| R-SQ | 0.96 | 0.95 | 0.95 |
| F EXCLUDED | 19.08 | 57.66 | 59.87 |
| N | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls as in table 9.

Our instruments are the local rainfall in the 1930s, the comparative advantage of the soil to grow cocoa and the distance to the railroad tracks. Except for specification (1), all instruments are significant and have the correct sign. Together, they are always strong. The exclusion restriction for the first two instruments holds because we control for the current rainfall on the pineapple farms and the soil suitability to grow pineapple. The exclusion restriction for the last instrument holds because in colonial times, the cocoa was transported via railway to the coast. Today, the railway is not used anymore for commercial purposes but pineapples are transported with trucks. Furthermore, the destiny of the pineapples are the processing companies in the Eastern Region, not the ports at the coast. A potential threat for the last instrument is the finding of Jedwab and Moradi (2015) that colonial infrastructure investments had a persistent welfare effect on their target regions. If the long-term income-effect of the railway has an effect on the self-efficacy of the farmers and the performance of the pineapple contracts, then the exclusion restriction would be violated. We argue that this is unlikely because we do not find an effect of the colonial cooperative success on price, which we would expect if those regions were significantly richer today.

### Table 12a. 2SLS First Stage – Self-Efficacy

| SPEC. | (1) | (2) |
|---|---|---|
| MODEL | 2SLS | 2SLS |
| DEP.VAR. | SELF-EFFICACY | SELF-EFFICACY |
| PREDICTED COOPS | 0.704** | 0.598** |
| | (0.263) | (0.207) |
| | [0.079] | [0.099] |
| SCHOOLS | -0.353 | -0.235* |
| | (0.182) | (0.114) |
| | [0.061] | [0.160] |
| ALL COVARIATES | YES | YES |
| DISTRICT FE | YES | YES |
| R-SQ | 0.181 | 0.269 |
| F EXCLUDED | 25.92 | 10.35 |
| N | 398 | 398 |

Notes: Predicted coops is a linear prediction from the analysis shown in table 11. Standard errors are bootstrapped to take into account the sequential estimation procedure (in round brackets) and regular standard errors are shown for comparison (in square brackets). They are all clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls as in table 9.

### Table 12b. 2SLS First Stage – Social Capital

| SPEC. | (1) | (2) |
|---|---|---|
| MODEL | 2SLS | 2SLS |
| DEP.VAR | SOCIAL CAPITAL | SOCIAL CAPITAL |
| PREDICTED COOPS | 0.767*** | 0.901*** |
| | (0.152) | (0.140) |
| | [0.061] | [0.057] |
| SCHOOLS | -1.028*** | -0.236** |
| | (0.151) | (0.0702) |
| | [0.034] | [0.084] |
| ALL COVARIATES | YES | YES |
| DISTRICT FE | YES | YES |
| R-SQ | 0.156 | 0.231 |
| F EXCLUDED | 694.79 | 20.65 |
| N | 398 | 398 |

Notes: Predicted coops is a linear prediction from the analysis shown in table 11. Standard errors are bootstrapped to take into account the sequential estimation procedure (in round brackets) and regular standard errors are shown for comparison (in square brackets). They are all clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls as in table 9.

Now we can use the predicted performance of the colonial cocoa cooperatives from our analysis and use it as instrument for self-efficacy in a subsequent 2SLS. If the colonial cocoa cooperatives really caused higher self-efficacy, our estimates should not differ much from the previous ones. However, if there is no causal effect from the performance of the colonial cocoa cooperatives on subsequent self-efficacy, then our final 2SLS model should produce significantly different estimates.

Tables 12a and 12b present the first stage of the procedure. It can be seen that our new instrument is strong and significant. To take into account our sequential estimation approach, we need to bootstrap our standard errors (first bracket under the coefficients). For comparison, we also show the regular standard errors (second bracket).

The estimated effect size is slightly larger than before we instrumented the colonial cooperatives. Table 12c presents the reduced form, which shows the same pattern. The estimated effect size is slightly increased in comparison to before but not qualitatively different. This suggests a causal effect of the performance of the colonial cocoa cooperatives instead of persistent self-efficacy differences through time.

**Table 12c. 2SLS Second Stage – The Cultural Determinants of Contract Farming Success**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | 2SLS | 2SLS | 2SLS | 2SLS |
| DEP.VAR | CONTRACT INCOME | CONTRACT INCOME | PERCENTAGE INCOME | PERCENTAGE INCOME |
| SELF-EFFICACY | 5.741***<br>(1.002)<br>(1.013) | 5.358***<br>(1.093)<br>(0.804) | 0.595***<br>(0.0135)<br>(0.104) | 0.442***<br>(0.0202)<br>(0.0411) |
| SOCIAL CAPITAL | 1.904***<br>(0.365)<br>(0.314) | 1.972***<br>(0.523)<br>(0.400) | 0.172***<br>(0.0109)<br>(0.0433) | 0.149***<br>(0.00107)<br>(0.0321) |
| ALL COVARIATES | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | -0.117 | -0.012 | -0.589 | 0.020 |
| N | 398 | 398 | 398 | 398 |

Notes: Predicted coops is a linear prediction from the analysis shown in table 11. Historic rainfall, soil suitability and distance to the railroad tracks were used to exogenize the successrate of the colonial cocoa cooperatives. Standard errors are bootstrapped to take into account the sequential estimation procedure (first brackets) and regular standard errors are shown for comparison (second brackets). They are all clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls as in table 9.

## VIII. Further Investigations

In this section, we investigate two more questions that are related to our previous findings. First, we estimate a two part model in which we estimate the participation in contract farming (part A) and the share of income generated through this channel (part B). Then, we consider the full income of the farmers.

The results of the two-part model are presented in table 13. It can be seen that self-efficacy and social capital both promote contract farming participation and the degree of participation. In contrast, the promotion of certification seems to only affect the former.

**Table 13. A Two Part Model for Contract Farming Participation and its Income Share**

| MODEL | PROBIT AND OLS | |
|---|---|---|
| PART | A | B |
| DEP.VAR | CONTRACT FARMING | PERCENTAGE |
| SELF-EFFICACY | 0.94*** | 0.40** |
| | (0.20) | (0.19) |
| SOCIAL CAPITAL | 0.22*** | 0.34* |
| | (0.023) | (0.20) |
| CERTIFIED | 0.58*** | 0.39 |
| | (0.18) | (0.47) |
| EDUCATION | 0.13 | 0.047 |
| | (0.14) | (0.21) |
| AGE | -0.18*** | 0.14*** |
| | (0.069) | (0.05) |
| LAND | 0.21*** | 0.18* |
| | (0.06) | (0.09) |
| CONTROLS | YES | YES |
| DISTRICT FE | YES | YES |
| R-SQ | 0.37 | |
| N | 398 | |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***).

Finally, we can have a look at the full income of the farmers.

To understand the determinants of farmers' full incomes, we require a measure of their off-farm incomes and other income sources, such as transfers and remittances. We do not have detailed income statements, but we have reported total income categories using five categories (less than 50 GHC per month on average, between 51 and 150, between 151 and 300, between 301 and 500, or more than 500). We use this dependent variable for the analysis shown in table 14. It is indicated that colonial cooperative performance has a positive effect and missionary schools have a negative effect and both self-efficacy and social capital have positive effects.

**Table 14. Full Income and its Determinants**

| SPEC. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| MODEL | OLS | OLS | 2SLS | 2SLS |
| DEP.VAR | INCOME CLASS | INCOME CLASS | INCOME CLASS | INCOME CLASS |
| COOPERATIVES | 0.134** | 0.154* | | |
| | (0.0465) | (0.0790) | | |
| SCHOOLS | -0.210** | -0.238* | | |
| | (0.083) | (0.119) | | |
| SELF-EFFICACY | | | 0.327*** | 0.263* |
| | | | (0.0591) | (0.141) |
| SOCIAL CAPITAL | | | 0.459*** | 0.225* |
| | | | (0.148) | (0.131) |
| CONTROLS | YES | YES | YES | YES |
| DISTRICT FE | YES | YES | YES | YES |
| R-SQ | 0.18 | 0.28 | 0.12 | 0.24 |
| F EXCLUDED SE | | | 62.45 | 52.59 |
| F EXCLUDED SC | | | 117.61 | 234.73 |
| N | 398 | 398 | 398 | 398 |

Notes: Standard errors are clustered at the district level. Significance levels are 10% (*), 5% (**), and 1% (***). Controls as in table 9.

## IX. Discussion and Conclusion

Our findings contribute to two lines of research. First of all, we show that culture affects the performance of contract farming, which adds to the existing work on contract farming, agricultural value chains, and rural development in Sub-Saharan Africa (Eaton and Shepherd 2001, Kirsten and Sartorius 2002, Kumar and Matsusaka 2009, Barrett et al. 2012, Bellemare 2012, Will 2013). Our two cultural traits are self-efficacy and social capital.

As discussed by Bandura (1997), self-efficacy beliefs “influence the course of action people choose to pursue, how much effort they put forth in given endeavors, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding [...] and their level of accomplishments they realize". Social capital on the other hand enables the farmers by providing support and help (Woolcock and Narayan 2000).

We also demonstrate that cultural differences between smallholder farmers can be explained with historical contexts, adding to the study of the historical origins of culture and its economic effects (Guiso et al. 2006, 2010, Tabellini 2010, Nunn 2012, Alesina and Giuliano 2013, Nunn 2013).

Our research also provokes new questions. First, we find Christian missionary schools to have a negative social capital effect which translates into an overall negative effect on the performance of contract farming. Previous research has usually found a positive effect of Christian missionary schools on various economic and political outcomes, mostly through the development of human capital (Woodberry 2004, Wantchekon et al. 2015). This raises the question of whether our finding is dependent on our focus on smallholder agriculture, where social capital is particularly important (Conley and Udry 2010, Pamuk et al. 2014, Wuepper et al. 2014), or whether there are other differences that produce this result.

Another question regards the speed of cultural change. Giavazzi et al. (2014) identify an interesting gap in the literature: Although culture is often defined in economics as the "customary beliefs and values that ethnic, religious, and social groups transmit *fairly unchanged* from generation to generation" (Guiso et al. 2006), there is actually a spectrum of different cultural traits, some of which persist fairly unchanged through time while others evolve and adapt rather rapidly (Giavazzi et al. 2014). We consider colonial cocoa cooperatives and Christian missionary schools and find that they shaped the evolution of self-efficacy and social capital, which currently affects contract farming performance. This could suggest that both cultural variables are only slowly adaptive, as Ghana's colonial era ended in 1957 and yet we still find its effect on self-efficacy and social capital present today.

An alternative view is suggested by the research conducted in India by Feigenberg et al. (2010) and in Colombia by Attanasio et al. (2009), who both find that social capital can be build up quickly by changing the frequency of social interactions, as well as in Ethiopia, by Meijerink et al. (2014) who find that social capital degraded quickly with the introduction of an institutional innovation, and Bernard et al. (2014) who find that all it takes to increase farmers' self-efficacy is to show them a motivating documentary with success stories of peers.

These apparently conflicting findings are possibly explained by the self-reinforcing nature of self-efficacy and social capital. For self-efficacy, e.g. Wuepper et al. (2016) show that low degrees of self-efficacy in the domain of agricultural investments lead to low returns on agricultural investments. Similarly, social capital is reinforcing as there is a larger benefit to cooperation in environments of high social capital than there is in an environment of low social capital, where cooperation might only be exploited (Tabellini 2008, Butler et al. 2009, Reuben et al. 2009). Thus, self-efficacy and social capital could be stable over generations because they were not affected by policies or other shocks for a long time. In contrast, effective policies, might not only have the potential to increase self-efficacy and social capital, but once they do so, they might even be supported from the same reinforcement mechanism that stabilized the two traits so persistently in the past.

As a cautionary note, it is important to note that self-efficacy and social capital can vary in magnitude (i.e., whether they are limited to simple situations or also reach into difficult ones), generality (how much they are correlated across different domains), and strength (how much it takes to reduce them). As Bandura (1997) emphasizes in the context of self-efficacy, observation of successful social peers and verbal persuasion can quickly increase self-efficacy, but such created self-efficacy is weaker than that which originates through personal successes. Similarly, social capital might be quickly increased through a policy but it needs time to grow strong. To build strong self-efficacy and social capital, time is needed, in addition to persistent effort.

For this reason, we recommend to not only focus on adults, but to specifically target children, perhaps when they are in school (Bandura 1997, Guiso et al. 2010). Recent examples of successful self-efficacy programs for adults include Jensen and Oster (2009) La Ferrara et al. (2012), and Bernard et al. (2014) who use electronic media to expose people to new ideas and worldviews. Specifically aimed at school children is the program designed by Krishnan and

Krutikova (2013), who designed a multifaceted program providing different sources of self-efficacy. Whereas the former examples are based on showing people social peers who demonstrate new behaviors, the latter is based on individual support, reflection, and encouragement. A recent successful example for how to build in a social capital policy in an existing program, Feigenberg et al. (2013) and Attanasio et al. (2009) show that social capital can sometimes be increased by simply increasing social interactions.

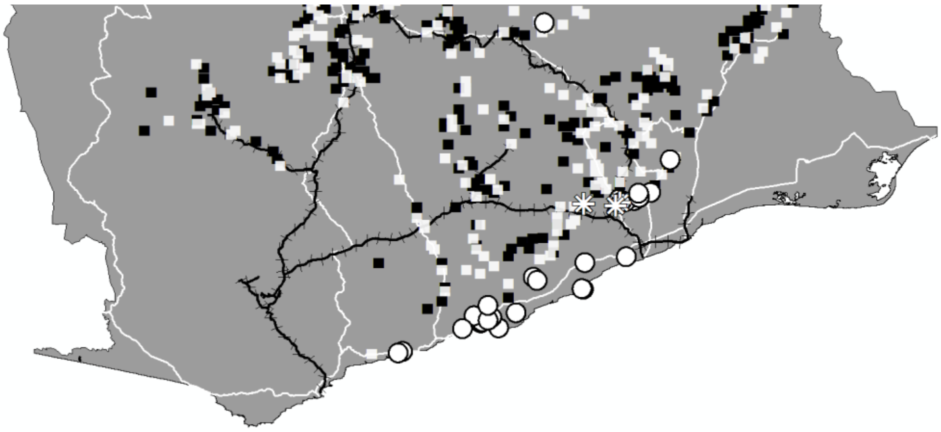
A simple way to target self-efficacy and social capital in Ghana would be to encourage the farmers more, give them more support and increase their social interactions during agricultural trainings. As the farmers in Ghana are frequently invited to agricultural trainings by various organizations and stakeholders, these trainings could easily be used to aim at self-efficacy and social capital, once their actual value is recognized.

It seems that addressing the beliefs of people can be as important as addressing other incentives and constraints (however, it is also not a substitute, as increased self-efficacy is strongly complementary to actual opportunity). A potential barrier to the implementation of policies that aim at self-efficacy and social capital is that most development initiatives have short planning horizons whereas culture needs time to change. Company managers and local stakeholders, however, might be the right change agents, as they can reap the medium- to long-term benefits.

In conclusion, we find that historical events explain cultural differences amongst Ghana's pineapple farmers and that such differences explain distinct performances of contract farming. Namely, colonial cocoa cooperatives and Christian missionary schools shaped the evolution of self-efficacy and social capital, which are both found to be important for contract farming. Ignoring this could lead to unrealistic expectations on the side of policy makers and business stakeholders and result in policies with unintended, adverse consequences. Example could include the provision of agricultural training that makes farmers doubt their abilities or any communication that encourages unrealistic expectations on the side of the farmers, subsequently leading to disappointment. Another example could be poorly designed contracts that tempt farmers not to stick to their contract, with an adverse effect on social capital.

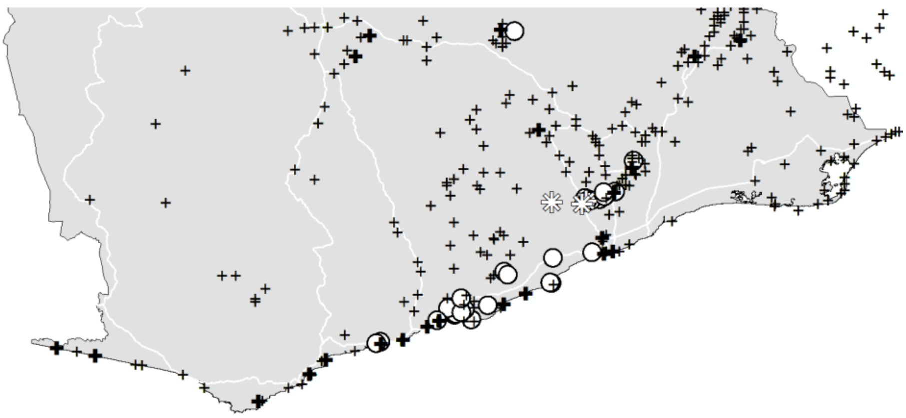
## ANNEX A

Colonial cocoa cooperatives in the 20th century



The map shows the south of Ghana. Black squares denote successful and white squares denote unsuccessful cocoa cooperatives. White circles denote locations of sampled farms and asterisks denote the companies. White lines denote roads and black lines denote colonial railroad tracks.

## ANNEX B



The map shows the south of Ghana. Bold crosses denote the missions, small crosses denote missionary schools, white circles denote locations of sampled farms, asterisks denote companies, and white lines denote roads.

# Chapter 2
# The Profitability of Investment Self-Efficacy: Agent-Based Modeling and Empirical Evidence from Rural Ghana

with Barbara Drosten[1]

**Abstract** Investment Self-Efficacy (SE) is the degree to which a decision maker believes to have the ability to increase her income through investing (money, time, effort). In addition to external incentives and constraints, SE is suggested to be a third, major determinant of investment behavior.

We begin with an agent-based model, simulating how historical environmental feedback might have caused different investment experiences and thereby investment SE. The model also demonstrates how such differences can cause current differences in income.

We test our model using empirical data from smallholder pineapple farmers in Ghana. Differences in investment SE are well explained by historical growing conditions for subsistence crops, such as cereals, roots and tubers, or tree crops. Historically determined investment SE subsequently explains significant income differences in Ghana. The channel is found to be a distinct investment behavior.

The content of this chapter is currently under peer-review with the Journal of Evolutionary Economics.

**Contributions** David Wuepper collected the data, created the agent based model, performed the econometric analysis and wrote most of the paper. Barbara Drosten contributed ideas to the survey, and provided content, especially to the introduction and the discussion section.

[1] Institute for International Politics, University of Hamburg

## 1. Introduction

Self-Efficacy (SE from here on) is the degree to which a person beliefs to have the capabilities to achieve her goals. This belief directly affects the choices a person makes as well as her performance (Maddux 1995, Pajares 2002) as it affects all intentional behavior. The more outcomes depend on choices and individual performance, the more SE matters (Bandura 1977, 1997, 2012, 2015). In contrast to general constructs such as locus of control, self-confidence and self-esteem, SE is domain specific, so that an individual can have high SE in performing for task A and low SE for task B (Bandura 1997). Individuals tend to choose goals that require what they believe to be good at, and they are more motivated to invest into achieving such goals.

SE is developed as a function of one´s own mastery experiences (pursuing and reaching an ambitious goal reassures the individual that she is competent), observed mastery experiences by social models (observing peers succeeding raises individual beliefs about own capabilities), verbal persuasion (encouragement has to be followed by mastery experiences to be robust though), and physiological arousal (situations causing stress or anxiety lower efficacy beliefs while positive emotions like excitement or joy increase it). An important source of SE are parents, who act as role models and convincingly communicate the qualities and potential of one´s capabilities (which connects SE to theories of identity, such as discussed by Akerlof and Kranton (2010), e.g.). To describe a person's SE, we can use 3 categories: magnitude (lower or higher levels of SE), strength (weaker or stronger beliefs decide about the degree of perseverance) and generality (how much SE correlates across domains).

As a simple example for the effect of SE, Weinberg et al. (1981) conducted an experiment in which individuals had to compete in an endurance task. Some individuals where told to have a weak competitor, the others where told to have a strong competitor (which increases and decreases SE, because here, one´s own competence is evaluated relative to the competitor). As suggested by SE theory (Bandura 1977, 2012), individuals with higher SE tried much harder than subjects with lower SE. Actually, all participants competed against professional athletes, so they all lost. In the second round, individuals with higher SE increased their effort to win, whereas individuals with lower SE decreased their efforts.

This is the basic mechanism of SE and it is plausibly important for a wide range of economic outcomes. Especially in more entrepreneurial occupations such as agriculture, SE could be just as important as external economic incentives and constraints (Wuepper and Sauer 2016, Wuepper et al. 2016).

In Ghana, as in most parts of Africa, the intensification of agriculture is a major challenge, and it is not well understood, why farmers are not investing more into making their operations more productive (Feder et al. 1985, Foster and Rosenzweig 2010, Dercon and Gollin 2014). Underlying most technology adoption models is the idea that farmers have different initial propensities to adopt an innovation, which can be explained with their characteristics (Zilberman et al. 2012). The task is to identify the sources of this heterogeneity. Interestingly, when it comes to subjective beliefs, adoption models are commonly Bayesian, in the sense that an initial belief is assumed, which is then logically updated in the light of new information. A major weak point is that there is no explanation where the initial belief comes from (Gilboa et al. 2012).

In the following, we propose a process of cultural evolution in response to environmental incentives to explain why some farmers inherit higher or lower SE from their ancestors. We model this process of culturally transmitted SE in an agent based model (in Netlogo). It is inspired by the mathematical models of Boyd and Richerson (1985), Bisin and Verdier (2001, 2010), and Galor and Özak (2014). The model predicts that environments which incentivized agricultural investments produced individuals with higher investment SE in the long-run, whereas environments with low returns on investments shaped a cultural evolution towards lower investment SE. As cultural traits tend to be very persistent, present SE beliefs may thus stem from forefathers and their environmentally fostered SE. The model also predicts that adaptation could be slow, so that changed incentives (e.g. the current context offers a high return on investment for everybody) may lead to inefficient investment levels for those farmers who have ancestors from a historically low return on investment context. We test our model with a sample of 400 pineapple farmers in the south of Ghana. The data is suitable for our study because Ghana's pineapple farmers are right on the interface between traditional and modern agriculture, giving us a helpful surrounding to study trajectories of development, and because we all the required data is available, on current economic behaviors and outcomes, on attitudes and beliefs, and about the farmers' ancestors and their lives.

For this study, we exploit the fact that the farmers have ancestors who were dependent on different subsistence farming systems and thus subject to historically different returns on investment. Nowadays~~,~~ they (would) all benefit from investing more intensive production (Suzuki et al. 2011, Kleemann and Abdulai 2013), so we can test whether farmers with roots in high return-on-investment regions achieve higher incomes than farmers with roots in low return-on-investment regions. Because of Malthusian dynamics, historically higher incomes mostly led to increases in family size~~s~~ and not to individual accumulation of capital. Thus, farmers differ in how much SE they have inherited and this is uncorrelated with how much capital they inherited. As a proxy test, we analyze whether farmers in the different regions inherited different farm sizes and do not find significant differences. We can also reject other causal channels, such as historical farming systems having systematically affected local prices, or off-farm income (as a proxy for economic development), social capital (as found in China by Talhelm et al. (2014) for other farming systems), or contract farming (because the companies are closer to one historical farming system than the other). Thus, we argue that historical farming systems can be used as instrumental variable~~s~~ to exogenize differences in observed SE. We find that investment SE is a strong predictor for actual investments and that such investments lead to large and significant income differences.

Recently, Gebrehiwot and van der Veen (2015) investigated the effect of SE differences on technology adoption in Ethiopia and also found a large and significant effect. As a caveat, their study is based on simple survey questions and they do not credibly control for unobserved heterogeneity. However, Bernard et al. (2014) conducted a randomized control trial (RCT), also in Ethiopia, and also find that increased SE increases investments. The only caveat of RCTs is that they commonly only identify short-term causal effects (e.g. aspirations), whereas many aspects of SE need a long time to develop (e.g. a strong and general belief in one's abilities). Thus, the contribution of our research is to identify the full, long-term effect of SE on agricultural investments and subsequent incomes; to identify the fundamental~~,~~ historical roots of observed SE differences~~, ;~~;and to better understand what SE actually means in economic terms and how it can be measured.

As a bonus, this is the first economic SE study from West Africa.

In the next section, we present our model (2). We then describe our analytical framework in section 3 and turn to our empirical analysis in section 4. We discuss our results in section 5 and conclude in section 6.

## 2. A Simple Agent Based Model

We develop our idea in a simple agent based model (ABM) in Netlogo. ABM is an alternative to equation based modelling (EBM), such as presented by Bisin and Verdier (2001) or Galor and Michalopoulos (2012). The main difference between the two approaches is that ABM are programmed and simulated and EBM are written and solved (Berger 2001, Janssen 2005, Farmer and Foley 2009). In many contexts, both approaches can be expected to give the same result but ABM can have advantages when it comes to complex interactions and emergent phenomena.

Our model begins with a population of historical subsistence farmers who can choose between two competing survival strategies: They can either try to minimize their costs and live off their natural endowments (endowment strategy) or they can try to increase their production and invest in~~to~~ their fields (investment strategy).

Farmers with high investment SE will naturally tend to choose the investment strategy. If such investments pay off, their SE increases and thus, their investments in the next period increase too.

If the investment does not pay off, they either increase their investments further (because high SE individuals tend to ascribe setbacks to their own insufficient effort which they can boost), or their SE decreases (if they ascribe failure to their ability or uncontrollable factors). This depends on the strength of their SE (e.g. farmers with success experiences have strong SE that is robust whereas inexperienced farmers without mastery experiences only have weak SE – see Bandura (1997) for ~~a~~ discussion).

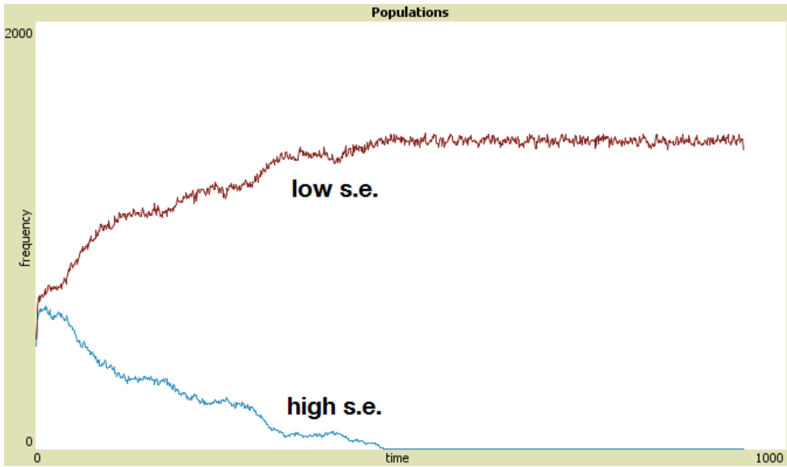
In contrast, farmers with low investment SE mostly choose the endowment strategy, without much consideration of investment opportunities. Their low SE tends to remain low as by forgoing risks of failure they also forgo the chance to succeed in mastering challenges and increase and strengthen their SE – until it may be raised externally. In the absence of targeted programs that combine verbal persuasion with vicarious experience and individual mastery experiences, ~~s~~uch external factors include only verbal and observing successful social peers. In these cases, farmers with low SE increase their SE and begin to invest too. If such investments pay off, their SE increases further, and they further increase their investments in the next period. If their investments do not pay off, their SE decreases again.

In our model, it is the environment that either rewards or punishes investments and thus determines whether the farmers develop higher or lower investment SE. Exogenous variables are the costs and benefits of investing, the probability of a random shock that negatively affects the return on investment (risk probability) and its severity (risk impact). Furthermore, we include institutions as a second source of feedback in addition to the natural environment. For simplicity, these institutions only affect the strength of the environmental feedback, so that environmental feedback might be weaker than without institutions.

We only present three exemplary cases for the cultural evolution of investment SE. In all figures, the X-axis is time and the Y-axis is the frequency of high and low SE in the population.

The situation depicted in figure 1 is an example for an environment in which investments are not sufficiently beneficial for the farmers. Thus, over time, the number of farmers with high SE dwindles, because beliefs are updated according to experiences.

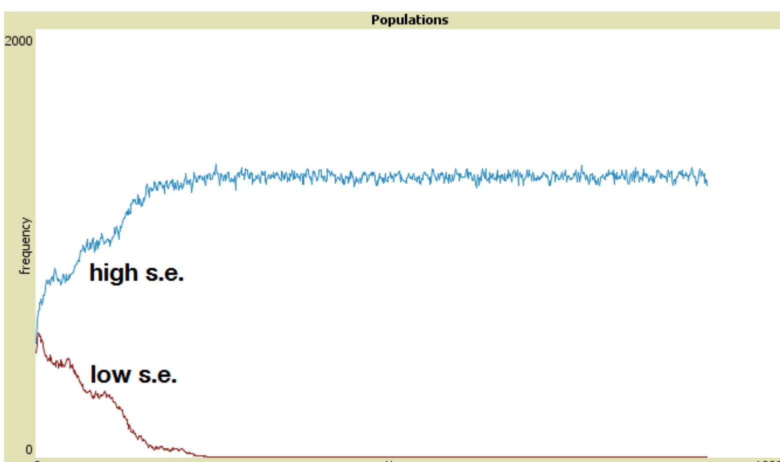
**Figure 1. The Evolution of SE (Context A)**




Cost=0.3 Benefit=0.4 Risk probability=0.1 Risk impact=0.5 Institutional persistence=0.5

In the situation depicted in figure 2, the return on investment is sufficiently high to let high SE evolve in this environment.

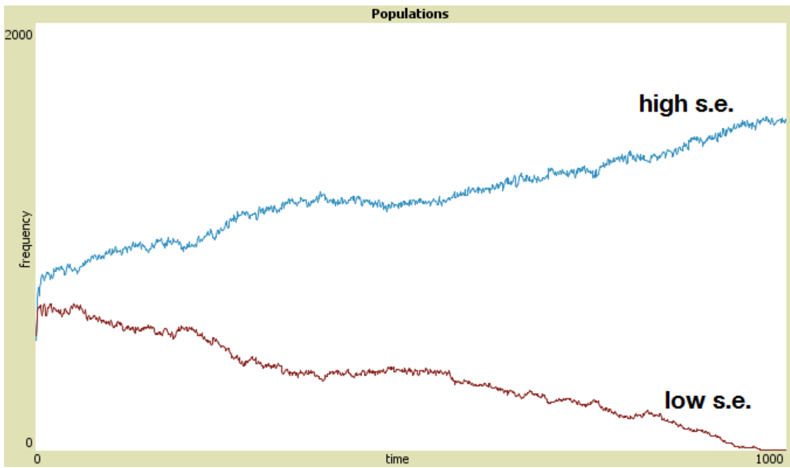
**Figure 2. The Evolution of SE (Context B)**




Cost=0.44 Benefit=0.66 Risk probability=0.2 Risk impact=0.4 Institutional persistence=0.2

In figure 3, it can be seen that institutions can slow down the cultural evolution of SE, if they are such that environmental feedbacks are dampened. The model ~~thus~~ predicts that we find higher investment SE in regions where historically, investments were sufficiently rewarded and we find lower SE in regions where historically, investments were sanctioned.

**Figure 3. The Evolution of SE (Context C)**




Cost=0.44 Benefit=0.66 Risk probability=0.2 Risk impact=0.4 Institutional persistence=0.8

The reason why historical environments affect current beliefs are learning costs. As discussed by Boyd et al. (2011) in general, and Nunn (2013) for economics, when learning is costly, it can be optimal for individuals to imitate others. Thus, instead of "re-inventing the wheel", individuals might decide to copy existing ones, without necessarily understanding all the details and without constant reappraisal of alternatives. For this reason, fundamental beliefs such as SE are usually found to be highly hereditary (Bandura 1997), but more culturally than genetically (Richerson and Boyd 2008).

It should be noted that cultural learning saves learning costs but at the cost of risking that mental models are outdated. Technological progress, environmental change, market developments and migration are a representative selection of factors that can make the knowledge of past generations ill-adapted to current circumstances. Re-considering figure 2, it can be seen that learning is not immediate, but there are several periods in which individuals hold low SE even though high SE is better adapted. In figure 3, where we allow institutions to dampen the environmental feedback, it takes almost the entire time horizon until adaptation is complete. Assuming that today most contexts offer profitable investment opportunities, the model predicts that we can explain a large share of income differences between individuals with their SE. Different levels of SE could also explain e.g. why many farmers in developing countries are reluctant to adopt technology (Feder et al. 1985) or to invest into highly profitable businesses (Udry and Anagol 2006).

In summary, our model suggests two testable hypotheses:

1) Historical environments explain current differences in investment SE and
2) Differences in investment SE have a causal income effect.

## 3. Empirical Framework

We begin with a naïve comparison of farmers with different incomes to see whether there is an empirical relationship between SE and income. To test the statistical significance of the observed relationship, we then proceed with baseline OLS regressions of the form

$$Y_{ij} = \alpha_j + \beta_1 SE_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \qquad (1)$$

where $Y_{ij}$ is the income of farmer $i$ in district $j$, $SE_{ij}$ are alternative measures for her SE , $X_{ij}$ is a vector of control variables, and $\alpha_j$ are district fixed effects. As we are analyzing a non-experimental, cross-sectional dataset, we are interested in the causes of SE:

$$SE_{ij} = \alpha_j + \beta_1 H_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \qquad (2)$$

where $H_{ij}$ are hypothesized historical causes of SE. To identify the causal effects of SE, we want to use $H_{ij}$ as instruments. To test the necessary exclusion restriction, we start by estimating 2SLS regression, in which we instrument alternative causal channels that could have been influenced by $H_{ij}$:

$$
\begin{aligned}
Y_{ij} &= \alpha_j + \beta_1 A_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \\
A_{ij} &= \alpha_j + \beta_1 H_{ij} + \beta_2 X_{ij} + \varepsilon_{ij}
\end{aligned} \tag{3}
$$

Where $A_{ij}$ are alternative channels through which $H_{ij}$ could affect income. We then proceed by estimating our 2SLS regression of interest:

$$
\begin{aligned}
Y_{ij} &= \alpha_j + \beta_1 Se_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \\
SE_{ij} &= \alpha_j + \beta_1 H_{ij} + \beta_2 X_{ij} + \varepsilon_{ij}
\end{aligned} \tag{4}
$$

The penultimate step is a 2SLS regression of investments on SE:

$$
\begin{aligned}
Inv_{ij} &= \alpha_j + \beta_1 Se_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \\
SE_{ij} &= \alpha_j + \beta_1 H_{ij} + \beta_2 X_{ij} + \varepsilon_{ij}
\end{aligned} \tag{5}
$$

The final step is a mediation analysis, to investigate how much investments mediate the SE effect.

## 4. Empirical Analysis

We begin this section with a description of our sampling strategy and our data. We then present the results of OLS regressions, followed by the results of 2 stages least squares regressions (2SLS), and finish with mediation analyses.

We representatively surveyed-400 pineapple farmers in the south of Ghana in 2013 (pineapple farming is only feasible in the south of Ghana as the north is too dry).

If the farmers want to participate in the exporting of fresh pineapple to the European Union, they need to have an export certification, which guarantees that certain production and quality standards are met (Kleemann and Abdulai 2013). Such certifications can be obtained from specialized organizations and are usually given to farm groups that are thus recorded and can be used for stratified random sampling. The procedure was as follows: First, the major pineapple growing areas were selected, and lists from groups of export-certified pineapple farmers were obtained. From these lists, farming groups were randomly selected and several farmers were interviewed reflecting the size of their group.

To cover non-certified farmers as well, extension agents and development agencies were asked to identify a representative sample of non-certified pineapple farmers for interviews. These non-certified farmers were sometimes in the same communities as the certified farmers and sometimes in adjacent communities.

The proportion of export-certified farmers in our sample is roughly 50%. It should ~~however~~ be noted, that many certified farmers do not make use of their certification and sell all their products at the local market, whereas some non-certified farmers sell their produce to companies for non-export usage. We present descriptive statistics about the farmers in table 3, but first we explain our variables (table 1) and especially, how we operationalize our SE measures (tables 1 and 2).

In table 1, the upper panel shows our SE measures. These measures focus on slightly different aspects of SE, and it will later be of interest to compare which are most relevant to the investment behavior of the farmers and their subsequent income.

The variable "nature" captures how much a farmer perceives nature to be providing for her, in contrast to being a potential that must be actively used. A farmer with high investment SE would tend to see nature as a potential. After a bad harvest, this farmer would increase her investments~~,~~ to increase it in the future. After a good harvest, she would increase her investment~~,~~ to further increase it (Bandura 2012). Such a farmer would also have a longer planning horizon. First, because this gives her additional opportunities to achieve more ambitious goals, and secondly, because a farmer with higher SE feels the responsibility to perform well (Fernandez et al. 2015). How much the farmer beliefs that her income is impacted by her abilities and choices are very direct measures for a farmer's SE (Bandura 2012). In contrast to the previous questions, an alternative measurement strategy is to openly ask the farmers what determined their incomes in the last two years.

**Table 1. Variable Descriptions**

| | |
|---|---|
| SE | Factor variable from the variables below |
| nature | Whether the farmer perceives nature to be a provider versus a potential (1-4) |
| planning | How long the farmer usually plans (from only for today to lifetime of children) |
| investing | How much the farmer believes that her income is determined by her investments |
| ability | How much the farmer beliefs her income to be determined by her abilities |
| determinant | Whether farmers report income determinants that are under their control (3), outside of their control (1), or ambiguous (2), from an open ended question. |
| social capital | Factor variable from the reported frequency of social event attendance, generalized trust, and number of people the farmer could borrow money from |
| off-farm income | Whether the farmer has off-farm income (1/0) |
| education | No. of years of formal schooling completed |
| age | Age of the farmer in years |
| household | Number of household members |
| gender | Whether the farmer is male (=1) or female (=2) |
| training | Whether the farmer received training from an international development agency |
| amount training | How often the farmer received the above trainings |
| Investments | Whether the farmer has *successfully* adopted an innovation in the last years |
| costs | all production related costs, excluding the opportunity costs of family labor |
| revenues | Quantity of sold pineapples times their price, on the local market and to companies |
| rain q | Reported, farm specific rainfall quantity |
| rain t | Reported, farm specific rainfall timing |
| soils | Reported, farm specific soil problems, defined as how much soils limit productivity |
| elevation | Calculated in GIS, in meters |
| topography | Calculated in GIS, as standard deviation of the elevation in meters |
| farm size | Area where pineapple is grown in hectares, including plots not currently used |
| prices | Average, local pineapple price |
| md2 variety | Whether the modern MD2 variety is grown |
| sc variety | Whether the Smooth Cayenne variety is grown |
| contract farming | Whether the farmer has a formal farming contract with a processor |
| tenure security | Reported safety of the plots |
| company distance | Calculated in GIS, as distance between farms and nearest company, in km |
| capital distance | Calculated in GIS, as distance between farms and Accra, in km |
| advantage cereals | Calculated in GIS , whether the biogeography is comparatively best for growing cereals |

This has the disadvantage, that answers must be coded into low, medium, and high SE by the researcher, which is risky for obvious reasons. However, it brings the important advantages that farmers are not influenced by the suggestions of the researcher but report what comes to their minds first. Furthermore, from the question, the farmers cannot tell what the research question is, so the associated biases can be avoided.

In table 2, we present a selection of typical answers and how they were coded.

**Table 2. Categorization of Mentioned Income Determinants**

| Under Control (3) | Ambiguous (2) | Outside Control (1) |
|---|---|---|
| agricultural practices | training | rainfall |
| seriousness | productivity | prices |
| regular weeding | yields | costs |
| learning | | diseases |

Notes: The table shows exemplary answers to the open question about the main income determinants in the last two years. The answers were translates into an index according to whether the mentioned determinants were under the control of the farmer (1), outside of her control (3), or whether the answer was not clearly within either category (2).

The vast majority of answers were farming related and quite clearly either under the control of the farmers or outside. If an answer describes a variable that is a choice of the farmer (such as agricultural practices, increased knowledge, improved attitudes and behaviors), then the variable "determinant" (for reported income determinant) was coded as a 3 (high SE). If the answer describes a purely exogenous variable (bad rainfall, low prices, high costs, etc.), then the variable "determinant" was coded as a 1 (low SE). Finally, if the answer was ambiguous or intermediate, it was coded as a 2. Examples for this last category are training (has the farmer chosen to participate in a training or does she attribute her low productivity to the fact that she is not provided with enough training), and yields (it is not clear which causes the farmer has in mind).

**Table 3. Comparing farmers in the lower, medium, and upper third of the farm-income range**

| income | low | | medium | | high | |
|---|---|---|---|---|---|---|
| statistic | mean | sd | mean | sd | mean | sd |
| SE | -.313 | .940 | .194 | .832 | .548 | 1.005 |
| determinant | -.155 | .961 | .017 | 1.019 | .454 | .953 |
| nature | -.238 | .855 | .087 | 1.051 | .552 | 1.073 |
| planning | -.268 | .892 | .156 | .991 | .551 | 1.057 |
| investing | -.188 | 1.126 | .159 | .755 | .296 | .848 |
| ability | -.157 | 1.102 | .188 | .778 | .148 | .950 |
| social capital | -.004 | 1.017 | -.030 | 1.087 | .068 | .761 |
| off-farm income | .186 | .390 | .196 | .399 | .343 | .478 |
| education | 2.641 | 1.205 | 2.663 | 1.103 | 2.955 | 1.307 |
| age | 44.698 | 11.987 | 43.975 | 9.330 | 43.761 | 9.055 |
| household | 5.842 | 2.766 | 5.819 | 2.571 | 6.432 | 3.129 |
| gender | 1.16 | .36 | 1.09 | .28 | 1.05 | .23 |
| training | .473 | .500 | .573 | .496 | .656 | .478 |
| amount training | 1.363 | .785 | 1.295 | .849 | 1.298 | .778 |
| investment | 3.24 | 2.71 | 4.09 | 2.50 | 5.02 | 1.93 |
| costs | 396.53 | 517.29 | 556.25 | 506.08 | 1200.60 | 984.22 |
| revenue | 380.59 | 351.44 | 1884.52 | 518.97 | 8280.62 | 5370.81 |
| rain q | 4.427 | 1.439 | 4.483 | 1.241 | 4.641 | 1.054 |
| rain t | 3.937 | 1.533 | 3.983 | 1.460 | 4.462 | 1.222 |
| soils | 1.674 | .746 | 1.655 | .820 | 1.567 | .820 |
| elevation | 88.825 | 66.451 | 86.013 | 62.311 | 72.117 | 39.887 |
| topography | 45.811 | 42.636 | 40.541 | 38.004 | 33.729 | 17.619 |
| farm size | 2.971 | 3.756 | 3.389 | 4.277 | 7.731 | 8.246 |
| prices | .431 | .187 | .373 | .134 | .404 | .096 |
| md2 variety | .200 | .401 | .368 | .484 | .552 | .501 |
| sc variety | .220 | .415 | .426 | .496 | .462 | .502 |
| contract farming | .349 | .477 | .344 | .477 | .641 | .483 |
| tenure security | 5.322 | 1.237 | 5.765 | .728 | 5.723 | 1.059 |
| company distance | 58615.14 | 35786.35 | 34021.36 | 30641.28 | 29376.49 | 31429.71 |
| capital distance | 65191.95 | 39899.9 | 40740.64 | 32765.05 | 37682.31 | 27894.74 |
| advantage cereals | .23 | .10 | .32 | .10 | .32 | .08 |

Notes: Variables in the upper panel are standardized (mean of zero and standard deviation of 1). The variables in the lower panel are unstandardized, except for social capital. In the analyses below, all variables are standardized.

The main explanatory variable in the following analyses is the factor variable SE, which reflects the commonality of all ~~the~~ individual SE measures. It is a factor variable from "determinant", "nature", "planning", "investing", and "ability".

Importantly, our SE indicators are only reported attitudes and beliefs – not outcomes. As an example, a SE measure is the attitude of the farmer towards investments, not how much she is currently investing. If we would include outcomes in our variable that we want to use to explain outcomes, we would not be the first to do this (see the discussion of Guiso et al. (2010) on the measurement of social capital) but we would be wrong.

Regarding our choice of control variables, we need to control for income determinants that are not endogenous to SE. Omitting variables that correlate with SE and income would create omitted variable bias. Including income determinants that are partially outcomes of SE would be "bad controls" (Angrist and Pischke 2008).

To get a feel for the data, table 3 presents descriptive statistics for three income groups (low, medium and high income). It can be seen that most SE measures increase with income but there are also other variables that increase with income. Financially better off farmers are more likely to grow the modern MD2 variety, or the Smooth Cayenne, which are both varieties that achieve higher prices on the market than other varieties. Such farmers also live closer to the capital and to processing companies and they have larger farms. Furthermore, a rugged topography is associated with ~~a~~ lower income, having been trained by an international development agency is associated with ~~a~~ higher income, as is the successful adoption an innovation in the last years and the extent of investments, such as in fertilizer.

To test whether the observed relationship between our different SE measures and income is statistically significant, we present the results of simple regressions in table 4. We include district fixed effects and control variables set 1 (which includes age, education, household size, gender, distances to Accra and to processors, rainfall quantity and timing, soils and topography). The standard errors are clustered at the farm group level. Table 4 shows that all SE measures are significantly and positively correlated with the income of the farmers, and the strongest relationship is found with the factor variable SE, which includes less measurement error than the other variables. From the individual measures, the variable "determinant" comes closest to the factor variable and "ability" is furthest away. The likely explanation is that "ability" is a rather suggestive question, whereas "determinant" does a better job in capturing the actual mental model of the farmers.

**Table 4. Baseline Results A**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| model | ols | ols | ols | ols | ols | ols |
| dep.var | income | income | income | income | income | income |
| determinant | 0.227*** | | | | | |
| | (0.0735) | | | | | |
| nature | | 0.184** | | | | |
| | | (0.0825) | | | | |
| planning | | | 0.216** | | | |
| | | | (0.0990) | | | |
| investing | | | | 0.132*** | | |
| | | | | (0.0407) | | |
| ability | | | | | 0.106* | |
| | | | | | (0.0614) | |
| SE | | | | | | 0.286*** |
| | | | | | | (0.0902) |
| controls | Set 1 | Set 1 | Set 1 | Set 1 | Set 1 | Set1 |
| district FE | yes | yes | yes | yes | yes | yes |
| R-sq | 0.19 | 0.17 | 0.18 | 0.16 | 0.15 | 0.20 |
| N | 398 | 376 | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

Because we always used the same set of control variables above, we need to establish that our estimates are robust to the inclusion of other control variables. In table 5, we start with no controls (spec. 1) and subsequently add more (spec. 2), and more (spec. 3) until we even include variables that might be endogenous (spec. 4). Set 1 is defined as before; set 2 additionally includes the farm size and local prices; and set 3 additionally includes the produced pineapple variety, whether the farmer has a farming contract with a processor, her tenure security, and whether she has off-farm income.

As table 5 indicates, the empirical relationship between SE and income is robust to variation in control variables and that even though the last two specifications include variables that are potentially already outcomes of higher investment SE (potentially "bad controls").

Having found this robust, empirical relationship, the question is whether there is an equally robust, causal effect of SE on the income of the farmers.

**Table 5. Baseline Results B**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | ols | ols | ols | ols |
| dep.var | income | income | income | income |
| SE | 0.269*** | 0.268*** | 0.231*** | 0.231*** |
| | (0.0744) | (0.0825) | (0.0771) | (0.0698) |
| controls | none | Set 1 | Set 2 | Set 3 |
| district FE | yes | yes | yes | yes |
| R-sq | 0.17 | 0.20 | 0.27 | 0.29 |
| N | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

To begin this analysis, table 6 presents the results from OLS regressions of SE on various explanatory variables. Afterwards, we estimate 2SLS regressions.

Table 6 indicates that especially regions with a historically, comparative advantage to grow cereals (instead of roots, tubers or tree crops) are now inhabited by farmers with higher investment SE. The data on regional suitability to grow different kinds of crops is publicly available in FAO's GAEZ database (http://www.fao.org/nr/gaez/en/).

**Table 6. Explaining SE (SE)**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | ols | ols | ols | ols | ols | ols |
| | SE | SE | SE | SE | SE | SE |
| adv. Cereals | 0.621*** | 0.575*** | 0.547*** | 0.621*** | 0.575*** | 0.547** |
| | (0.110) | (0.124) | (0.144) | (0.121) | (0.117) | (0.222) |
| gender | -0.0241 | -0.0427 | 0.00501 | -0.0241 | -0.0427 | 0.00501 |
| | (0.184) | (0.177) | (0.188) | (0.185) | (0.196) | (0.206) |
| education | -0.0965* | -0.0946* | -0.0929* | -0.0965** | -0.0946** | -0.0929 |
| | (0.0503) | (0.0508) | (0.0510) | (0.0416) | (0.0443) | (0.0580) |
| age | 0.0735 | 0.0711 | 0.0825 | 0.0735*** | 0.0711** | 0.0825*** |
| | (0.0441) | (0.0461) | (0.0508) | (0.0269) | (0.0286) | (0.0246) |
| rain q | -0.143* | -0.148* | -0.145* | -0.143** | -0.148** | -0.145*** |
| | (0.0789) | (0.0772) | (0.0764) | (0.0640) | (0.0609) | (0.0513) |
| rain t | 0.230** | 0.213** | 0.233** | 0.230* | 0.213* | 0.233** |
| | (0.101) | (0.0944) | (0.0905) | (0.136) | (0.129) | (0.118) |
| soils | -0.0433 | -0.0497 | -0.0324 | -0.0433 | -0.0497 | -0.0324 |
| | (0.0433) | (0.0454) | (0.0405) | (0.0362) | (0.0332) | (0.0386) |
| elevation | -0.229 | -0.205 | -0.217 | -0.229* | -0.205* | -0.217* |
| | (0.170) | (0.171) | (0.169) | (0.128) | (0.115) | (0.112) |
| topography | -0.438* | -0.383* | -0.453* | -0.438*** | -0.383*** | -0.453*** |
| | (0.245) | (0.213) | (0.230) | (0.0868) | (0.0743) | (0.102) |
| training | | 0.0727 | 0.0520 | | 0.0727 | 0.0520 |
| | | (0.0798) | (0.0827) | | (0.0734) | (0.0859) |
| amount training | | -0.116 | -0.118 | | -0.116** | -0.118*** |
| | | (0.0839) | (0.0758) | | (0.0498) | (0.0426) |
| household | | | -0.0244 | | | -0.0244 |
| | | | (0.0369) | | | (0.0421) |
| capital distance | | | 0.540 | | | 0.540 |
| | | | (0.834) | | | (1.011) |
| company distance | | | -0.478 | | | -0.478 |
| | | | (0.777) | | | (0.848) |
| farm size | | | 0.101** | | | 0.101*** |
| | | | (0.0423) | | | (0.0335) |
| price | | | -0.00320 | | | -0.00320 |
| | | | (0.0430) | | | (0.0256) |
| tenure security | | | 0.186*** | | | 0.186*** |
| | | | (0.0417) | | | (0.0509) |
| district FE | yes | yes | yes | yes | yes | yes |
| R-sq | 0.36 | 0.36 | 0.41 | 0.36 | 0.36 | 0.41 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are multi-dimensionally clustered at the district and the community level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

Most current context variables, such as whether the farmer has been trained by an international development agency, the local pineapple price, or the distances to Accra or the processing companies, do not have significant explanatory power.

The amount of rainfall and education are estimated to be slightly negatively correlated with SE. However, the timing of ~~the~~ rain and a less rugged terrain are positively correlated. Also significantly, positively correlated are farm size and tenure security, but these variables are plausibly endogenous. Strikingly, the strongest correlation is found between the historical advantages to grow certain kinds of crops and SE. Below, we will exploit this relationship to exogenize SE.

The reason why ~~a~~ historical advantage to grow cereals caused a persistent change in the evolution of SE is its effect on the historical choice of farming systems (Michalopoulos et al. 2016). Biogeographic circumstances strongly affected which production systems Ghana's pre-colonial communities chose.

Along the coast of the Central Region, i.e., the farmer found good conditions to grow tubers and roots. Roots and tubers have several advantages. Amongst others, they can be grown and harvested all year long, so they require less planning than other crops, and they are robust and flexible, so mistakes and production constraints are less severe. Their most prominent characteristic is their very low requirement for inputs (Rees et al. 2012). A very different kind of crops are cereals. Cereals are less robust and flexible, and require far more investments. On the other hand, they respond strongly to investments and their productivity can be greatly increased with the right agricultural practices at the right time (Heisey and Mwangi 1996). The region in the south of Ghana that had a comparative advantage to grow cereals is located in the Savanna zone north of Accra. Even further north, biogeographic circumstances favored tree crops, ~~whose~~ the production of which is closer to that of roots and tubers than to that of cereal crops. The explanation why such a historic variable can persist to affect current farmer behavior is cultural evolution as discussed by Nunn (2012, 2013), Henrich et al. (2001), Boyd et al. (2011) and Henrich et al. (2008). The basic idea is that individuals can save learning costs by imitating their parents and other social peers. Even without understanding why, a farmer might behave well adapted in her environment if she chooses e.g. similar investment levels as others in her community have always done. This strategy, however, works better in environments that do not change too much, as past behavior of one's ancestors has evolved to fit to their context and not to the current one. Thus, migration and technical change are two classic examples why culture can become "outdated" (Richerson and Boyd 2008)

**Table 7. Direct Economic Effects of Historic Farming Systems**

| . | (1) | (2) |
|---|---|---|
| | OLS | OLS |
| dep.var. | land endowment | pineapple price |
| advantage cereals | 0.142*** | 0.105 |
| | (0.0349) | (0.189) |
| advantage tubers | 0.110*** | 0.0471 |
| | (0.0369) | (0.107) |
| controls | yes | yes |
| district FE | yes | yes |
| R-sq | 0.13 | 0.92 |
| N | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

An important step is to proof that historical biogeographic differences did not affect other causal channels that currently affect the income of the farmers. Because cereals, roots, and tubers have been (and continue to be) mainly subsistence crops and because more food strongly increased family sizes, it is unlikely that they created economically significant differences in material endowments between the regions. Regressing the land endowment of the pineapple farmers and their local pineapple price on whether their ancestors mostly farmed cereals or roots and tubers show, that differences in historic farming systems did not create important differences in material endowments between the farmers. We control for district fixed effects as well as biogeographic variables, age and education.

A second set of tests is presented in table 8, which shows the results of three 2SLS specifications, in which income is the dependent variable and we use the historic ~~,~~ comparative advantage to grow cereals as instrument for social capital (spec. 1), off-farm income (spec. 2),and contract farming participation (spec.3). We control for age, education, household size, gender, distances to Accra and to processors, rainfall quantity and timing, soils and topography (control variables set 1). We also include district fixed effects and cluster the standard errors at the farmer group level.

**Table 8. Falsification Test: Instrumenting Confounding Variables**

| | (1)<br>2SLS | (2)<br>2SLS | (3)<br>2SLS |
|---|---|---|---|
| 2ND STAGE | income | income | income |
| social capital | -15.08<br>(45.34) | | |
| off-farm income | | 2.717**<br>(1.332) | |
| contract farming | | | -27.85<br>(97.65) |
| 1ST STAGE | social capital | off-farm income | contract farming |
| adv. cereals | -0.0417<br>(0.130) | 0.231<br>(0.140) | -0.0226<br>(0.0797) |
| controls | set 1 | set 1 | set 1 |
| district FE | yes | yes | yes |
| R-sq 2 | -181.927 | -5.765 | -130.628 |
| R-sq 1 | 0.203 | 0.127 | 0.310 |
| F excluded | 0.10 | 2.74 | 0.08 |
| N | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

Notably, in table 8, no first stage indicates any meaningful causal effect. The F-statistic is below 3 in all three specifications and the advantage to grow cereals is insignificant. This tells us that historic farming systems did not affect current levels of social capital, the probability to obtain off-farm income, or the participation in contract farming.

This suggests that historical production systems did not significantly (in the economic sense of the word) affect the financial resources of the farmers (which would help them to participate in contract farming, and neither did it create local employment possibilities for the farmers (which would give them access to off-farm income) , or increase their ability to work together (as captured with their social capital).

Thus, we are confident that a biogeographic advantage to grow cereals only affects the investment SE of the farmers and no other income determinant and thus, it is a feasible instrument for the SE of the farmers. It is also a better instrument than historic production systems, e.g. cereal farming, or roots farming. The reason is that historical production systems could well be endogenous. Societies with higher investment SE could have chosen production systems that rewarded investments , and societies with lower SE could have chosen production systems that require less investment. The regional suitability to grow different crops, in contrast, affected which crops were grown and could not be changed by the farmers.

In table 9, we present the results from 2SLS regressions of income on SE and different control variables. The control variables set are defined as for table 5. As before, we start with no control variables, include a few, clearly exogenous variables, and then increase the number of control variables until we also include potentially endogenous ones. We always include district fixed effects and cluster the standard errors at the farmer group level. Looking at the first stage, it can be seen that we always find a significant effect of the comparative advantage to grow cereals on the SE of the farmers. The F-test is always 20, so we clearly have a strong instrument here. Looking at the second stage, we find a significantly larger, causal effect for the farmers' SE on their income.

Specifications (3) and (4) are not reliable because of potentially endogenous control variables. However, specification (2), which is theoretically the most sound specification, indicates that indeed the causal effect is larger than the mere correlation in table 5 could suggest (it should be noted that we use a factor variable reflecting a cultural trait. Thus, the magnitude of the estimated causal effect is not trivially interpretable).

**Table 9. The Causal Effect of SE (SE) on Income**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | 2sls | 2sls | 2sls | 2sls |
| 2ND STAGE | income | income | income | income |
| SE | 0.397***<br>(0.126) | 0.369*<br>(0.210) | 0.796***<br>(0.303) | 0.833**<br>(0.347) |
| 1ST STAGE | SE | SE | SE | SE |
| adv. Cereals | 0.841***<br>(0.158) | 0.619***<br>(0.114) | 0.658***<br>(0.147) | 0.593***<br>(0.127) |
| controls | none | set 1 | set 2 | set 3 |
| district FE | yes | yes | yes | yes |
| R-sq 2 | 0.16 | 0.18 | 0.07 | 0.09 |
| R-sq 1 | 0.25 | 0.36 | 0.37 | 0.43 |
| F excluded | 28.38 | 29.64 | 20.10 | 21.67 |
| N | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

For obvious reasons, finding feasible measures for a farmer's SE is inherently difficult. To investigate which of our individual measures do a better job than others, we present in table 10 the results from 2SLS regressions of income on the individual SE measures, which we instrument as before.

Using a factor variable, instrumented with a theory provided instrument is clearly the preferred approach. However, testing our individual measures can inform us, which individual measures to use in the future, and which to substitute or change.

The first stage in table 10 indicates that our instrument is weak for the open question about past income determinants ("determinant") and for the question about the importance of one's ability as an income determinant ("ability"). The first stage for the question whether nature is rather a provider or rather a potential is also somewhat weak, but feasible. Looking at the second stage, all measures except "ability" are significant and positive. Notably, the magnitude of the estimated individual measures effects is clearly larger than that of the factor variable.

**Table 10. Evaluating SE Measures**

| model | (1)<br>2sls | (2)<br>2sls | (3)<br>2sls | (4)<br>2sls | (5)<br>2sls |
|---|---|---|---|---|---|
| 2ND STAGE | income | income | income | income | income |
| determinant | 2.115**<br>(0.836) | | | | |
| nature | | 1.148***<br>(0.359) | | | |
| planning | | | 1.056***<br>(0.298) | | |
| investing | | | | 1.011**<br>(0.472) | |
| ability | | | | | 2.333<br>(1.567) |
| 1ST STAGE | determinant | nature | planning | investing | ability |
| adv. Cereals | 0.297**<br>(0.134) | 0.570***<br>(0.204) | 0.595***<br>(0.189) | 0.621***<br>(0.171) | 0.269*<br>(0.136) |
| controls | set 1 | set 1 | set 1 | set 1 | set 1 |
| district FE | yes | yes | yes | yes | yes |
| R-sq 2 | -2.769 | -0.474 | -0.314 | -0.520 | -3.797 |
| R-sq 1 | 0.19 | 0.30 | 0.33 | 0.17 | 0.23 |
| F excluded | 4.90 | 7.76 | 9.91 | 13.25 | 3.94 |
| N | 398 | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

In conclusion, asking farmers how much they perceive their abilities to determine their income does not seem to be a good way to capture their SE. Plausibly, the question is too direct and farmers cannot answer this question accurately and objectively.

Above, we find the following causal chains:

Historical Production Systems => Investment SE => Income

The obvious gap in this analysis is the connection between investment SE and income. What we need is a final analysis showing that investment SE affects incomes through investments. We begin by showing that SE indeed affects investments. In table 11, we present four specifications, with two different measures for investments.

**Table 11. The Effect of SE on investments**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | ols | ols | ols | ols |
| dep.var | investment | investment | investment costs | investment costs |
| SE | 0.523*** | 0.539*** | 0.230*** | 0.239*** |
| | (0.0837) | (0.0768) | (0.0554) | (0.0575) |
| amount training | | 0.125*** | | 0.148*** |
| | | (0.0406) | | (0.0460) |
| credit | | 0.104** | | 0.237** |
| | | (0.0441) | | (0.0994) |
| controls | set 1 | set 1 | set 1 | set 1 |
| district FE | yes | yes | yes | yes |
| R-sq | 0.30 | 0.33 | 0.16 | 0.23 |
| N | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are multi-dimensionally clustered at the district and the community level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

In specifications (1) and (2), the dependent variable is whether the farmer has successfully adopted an innovation in recent years. In specifications (3) and (4), the dependent variable are the annual production costs of the farmer. The production costs are monetary, so unpaid labor is omitted. However, the farmers use mostly paid labor for planting and harvesting pineapples, and in general, monetary costs are strongly correlated with general investments, as low monetary costs usually indicate a lack of productive inputs other than land, labor, and planting material (Kleemann and Abdulai 2013).

It can be seen that SE is always a significant investment determinant, whether defined as having invested into an innovation or the height of production costs. In specifications (2) and (4), we include the potentially endogenous control variables "amount of training" (from international development agencies) and "credit", which captures whether the farmer currently has a credit.

Farmers with higher SE could be more likely to participate in training and seek-a credit because they believe in their ability to exert influence. However, the estimated effect of SE is (a) larger when these variables are included (possibly, trainings are more often offered to farmers with low SE) and (b), the estimated effect of SE is larger than those two variables, which are amongst the most prominent explanations for farm investments.

**Table 12. Mediation Analysis for the SE Effect**

| spec. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | mediation | mediation | mediation | mediation |
| dep.var. | income | income | income | income |
| SE | 0.212*** | 0.185*** | 0.184*** | 0.155*** |
| | (0.0567) | (0.0632) | (0.0478) | (0.0533) |
| investments | 0.116** | 0.113** | | |
| | (0.0540) | (0.0555) | | |
| costs | | | 0.370*** | 0.376*** |
| | | | (0.0457) | (0.0464) |
| ACME | .056*** | .060*** | .085*** | .089*** |
| Direct Effect | .211*** | .185*** | .183*** | .155*** |
| Total Effect | .267*** | .246*** | .268*** | .245*** |
| % of Tot Eff mediated | 21% | 24% | 32% | 36% |
| controls | none | Set 1 | none | Set 1 |
| district FE | yes | yes | yes | yes |
| R-sq | 0.18 | 0.20 | 0.29 | 0.31 |
| N | 393 | 393 | 393 | 393 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

We continue with table 12, in which we show the results of a mediation analysis. In specifications (1) and (2) we use a dummy mediation variable, which captures whether the farmers report to have successfully adopted an innovation in recent years. This is indicated to mediate the investment SE effect by 21 – 24%. In specifications (3) and (4) we use the average costs of investments over the last 2 growing periods as mediation variable. This is indicated to mediate between 32 and 36% of the investment SE effect. These figures are clearly below 100% but this is consistent with SE theory, because much of the investments is unobserved in form of individual effort, time spent working and work~~ing~~ intensity.

**Table 13. Mediation Analysis Including the Effect of Risk**

| spec. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | mediation | mediation | mediation | mediation |
| dep.var. | income | income | income | income |
| SE | 0.225*** | 0.202*** | 0.175*** | 0.149** |
| | (0.0666) | (0.0729) | (0.0590) | (0.0641) |
| investment | 0.117** | 0.112* | | |
| | (0.0569) | (0.0585) | | |
| costs | | | 0.379*** | 0.384*** |
| | | | (0.0465) | (0.0472) |
| ACME | .048*** | .050*** | .098*** | .104*** |
| Direct Effect | .223*** | .201*** | .173*** | .148*** |
| Total Effect | .271*** | .252*** | .272*** | .253*** |
| % of Tot Eff mediated | 18% | 20% | 36% | 41% |
| controls | set 6 | set 7 | set 6 | set 7 |
| RISK CONTROLS | yes | yes | yes | yes |
| district FE | yes | yes | yes | yes |
| R-sq | 0.19 | 0.21 | 0.30 | 0.32 |
| N | 393 | 393 | 393 | 393 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at farmer group level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

Spending money and adopting technology is only a part (indicated to be between 20 and 40%) of the investments. A final robustness check is to include various risk measures in our analysis (table 13). Having different risk perceptions and attitudes is not independent from SE (basically SE leads to domain specific (e.g. investments) risk perceptions and attitudes). However, including general risk perceptions and attitudes might enable us to estimate SE more precisely than before. Thus, we include perceived financial riskiness of pineapple farming, perceived ability to take on financial risk, perceived pay-off to risk taking, and the choice in a small experiment with differently risky choices. Neither of these variables is significant once we include investment SE and the estimated mediation effect does not substantially change.

## 5. Discussion

We developed a simple agent-based model to understand how environmental feedback might have shaped the evolution of investment SE, and how this could affect present incomes. Our empirical data from Ghana's pineapple farmers is consistent with the idea that historical suitability for different subsistence crops created distinct experiences with agricultural investments, which is why today, farmers in different regions of Ghana have distinct degrees of SE. This causes them to invest differently in their production and obtain different incomes. We analyze Ghana's pineapple farmers on income, SE, historical environments, and confounding factors. We have reason to believe that one would obtain the same pattern with a global dataset, or in different regions. This, however, is left for future research.

It is important to note that SE is subject to several feedbacks. One such feedback is that low SE is a causal factor for poverty (Bernard et al. 2011, Bernard et al. 2014) and poverty causes low SE (Dercon and Krishnan 2009, Dercon and Singh 2013). Another one is that low SE reduces ambitions, which are important for economic development as well as SE development (Bandura 2012, Flechtner 2014, Pasquier-Doumer and Brandon 2015, Dalton et al. 2016). However, there are multiple successful initiatives that demonstrate how SE can be increased. Krishnan and Krutikova (2013) conducted an innovative program in an Indian slum, where they significantly increased the SE of poor high school students by giving them challenging tasks, while supporting them to succeed in them, and by providing a mentor, with whom they discussed their ambitions.

Such programs are comparatively expensive but also effective. Far less expensive are media based approaches, such as the one chosen by Bernard et al. (2014). Showing poor smallholder farmers in Ethiopia a documentary about success stories of social peers already raised SE. Similarly positive experiences with media based solutions have been made by Jensen and Oster (2009) and La Ferrara et al. (2012) in other contexts. (Bandura himself showed in his famous Bobo doll experiments in the 1960s the strong effect of media, identification figures and imitating behavior). It should, however, be noted that the strength of SE beliefs differs depending on their source. SE strength describes how much adversity is needed to reduce it. If a person's SE has been raised ( by verbal persuasion or observing the success of social peers) followed by one´s own mastery experience of success then her SE belief is probably robust–enough to survive failures and adversity; if the personal mastery experience is missing, it may not hold up. This is why a good training should combine SE targeted actions with external constraints targeted actions, to raise aspirations and increase the chance of mastery experiences as much as possible.

To use SE as a poverty mitigation tool, policies must be designed to both raise and strengthen SE. It should be noted that SE only affects the performance of an individual. When outcomes do not depend much on individually influencable factors like performance but predominantly on other constraints, like capital or infrastructure or property rights, there is not much of a role for SE. If, however, constraints make the achievement of ambitious goals more difficult, but not impossible, SE enhanced aspirations and performance can make a big difference (Pritchett and Kapoor 2009).

Wuepper and Sauer (2016) find that human and social capital can partially compensate for low levels of inherited SE. We expect similar effects for good institutions and individual leadership, which could be investigated in the future.

SE also deserves further academic attention for another reason. As we find, SE is a fundamental behavioral determinant that is shaped by long-term environmental factors and then by multiple feedbacks, e.g. own experiences or the experiences of social peers. It might explain a wide spectrum of observed behavioral factors, such as domain specific risk attitudes and perceptions (Weber et al. 2002, Nicholson et al. 2005), time preferences (Galor and Özak 2014), and preference for entrepreneurship (Galor and Michalopoulos 2012, Kautonen et al. 2013), to name a few examples. For technology adoption models, it should be noted that SE could both explain the outcome (whether a technology is adopted, whether it is dis-adopted, the extent of adoption, its profitability) and many explanatory variables (education, credit access, risk preferences, …), so there is a risk of omitted variable bias when SE is not included in the model.

## 6. Conclusion

We find that without policy intervention, investment SE is largely inherited from one's ancestors, whose SE was adapted to historical environmental feedback to agricultural investments. Without updating, inherited SE is not everywhere adapted to current incentives and constraints. This causes a forgoing of chances for a share of the population. We find that farmers with low SE achieve significantly lower incomes than their high SE peers. We also expect that SE is involved in many more phenomena than currently realized but this requires further research. Since we tested our hypotheses only amongst the pineapple farmers of Ghana, the next logical step is to investigate the role of SE for other crops and sectors and other regions. Given our results, SE seems a promising lever for poverty mitigation.

# Chapter 3

# Self-Efficacy or Farming Skills: What Matters more for the Adaptive Capacity of Ghana's Pineapple Farmers?

with David Zilberman[3] and Johannes Sauer[4]

**Abstract** Self-efficacy is the subjective belief of an individual to have the ability to achieve a goal. We test whether this belief is actually more important than a farmer's objective skills in determining her adaptive capacity to decreasing rainfall. Using data of 400 farmers between 2009 and 2013, we investigate the probability to adopt a climate smart technology in response to drought. We find that self-efficacy is indeed more important than objective farming skills to explain behavior and this is robust to different specifications and 2SLS estimation. Finally, we find evidence that farmers with higher self-efficacy suffer less income loss from adverse weather than farmers with lower self-efficacy. As a cautionary note, self-efficacy and objective ability are complementary to each other, so there is no point in focusing on the self-efficacy of an individual who is bindingly constraint just as it can be inefficient to aim at behavioral changes without considering self-efficacy.

This chapter is currently under peer review with the Journal of Agricultural Economics

**Contributions** David Wuepper and David Zilberman came up with the research question while David Wuepper's research visit at the University of California Berkeley in 2015. David Wuepper also prepared and analyzed the data, and wrote the article. David Zilberman and Johannes Sauer contributed to the research with their ideas, discussion, and suggestions. This chapter also benefited from the feedback of several attendees of the 2016 CSAE conference at the University of Oxford.

[3] University of California, Berkeley, Department of Agriculture and Resource Economics
Technical University Munich, Agricultural Production and Resource Economics

## 1. Introduction

Economists have long realized that individuals with apparently similar incentives and constraints show different technology adoption behaviors (David 1975, Feder et al. 1985). Explaining the individual thresholds to technology adoption continues to be an active area of investigation, especially whether adoption heterogeneity is caused by external (e.g. infrastructure, financial access) or internal (e.g. psychology, human capital ) factors (Foster and Rosenzweig 2010, Duflo et al. 2011, Suri 2011). A growing research area is individual climate change adaptation and how it differs from other areas of technology adoption (Zilberman et al. 2012, Di Falco 2014, Arslan et al. 2015). In the following, we analyze the decision of smallholder pineapple farmers in Ghana to adopt a climate smart innovation in response to experiencing decreasing rainfall. The adaptation to changing circumstances requires a particular set of skills skill, which is differently called in different research communities. In agricultural economics, Theodore W. Schultz (1975, 1980) coined the term Allocative Ability to describe the ability of an individual to effectively respond to change. In the climate change literature, the term Adaptive Capacity is used, which describes the individual ability to effectively adapt to climate change (Yohe and Tol 2002, Smit and Pilifosova 2003, Smit and Wandel 2006). Initially, both concepts were free of psychology and culture but recent empirical evidence suggests that especially self-efficacy is an important explanation (Grothmann and Patt 2005, van Duinen et al. 2014, Gebrehiwot and van der Veen 2015). Self-efficacy captures how much an individual beliefs to have the necessary ability to achieve a certain goal (Bandura 1977, 1997, 2012). This belief is domain specific, so in contrast to general personality traits such as self-confidence, locus of control, or self-esteem, an individual can have high self-efficacy in some domains and low self-efficacy in others. Research by Wuepper and Drosten (2016) and Wuepper and Sauer (2016) suggests that self-efficacy is to a large extend culturally inherited and that it only evolves slowly. What is inherited from our ancestors can be thought of as Bayesian priors (Bisin and Verdier 2010). How these priors are updated to posterior beliefs depends on four sources, which are own mastery experiences (pursuing and reaching an ambitious goal, which tell the individual that she is competent), observed mastery experiences (seeing a social peer to success conveys information to the individual that she probably can do it to), persuasion (being told to be competent is likely to be less effective than the former two sources but can nevertheless raise one's self-efficacy, if the source of the information is convincing), and finally, emotions (negative emotions lower self-efficacy while positive emotions increase it). For extensive non-economics overviews of self-efficacy, the reader is referred to the books of Bandura (1995, 1997) and Schwarzer (2014).

Our main contribution is that we quantify the effect of self-efficacy on the adaptive capacity of farmers in West-Africa – after it has been found significant in northern Europe (van Duinen et al. 2014) and East Africa (Gebrehiwot and van der Veen 2015) and we explore its income effect. Methodologically, we are the first to address the issue of endogeneity in this context.

In the next section (2), we briefly discuss some background literature and how self-efficacy can be measured. We then turn to our data (3) and how we operationalize our variables. In section 4, we present empirical results on the effect of self-efficacy on technology adoption behavior and in section 5, we explore its income effect. Section 6 concludes with a discussion of research and policy implications.

## 2. The Effects of Self-Efficacy and their Measurement

We investigate the question whether objective farming skills or subjective self-efficacy beliefs are more important for the adaptive capacity of smallholder pineapple farmers in Ghana. This is motivated by the fact that until recently, adaptive capacity, or allocative ability, where entirely explained without psychological or cultural factors, and objective farming skills are a likely alternative explanation to self-efficacy once we control for external constraints such as credit, infrastructure, and market factors. However, as we briefly discuss below, this should not be understood as self-efficacy and farming skills being good substitutes for each other. After a brief summary of some empirical literature, we clarify this point and why it matters.

### 2.1. Survey based Evidence

Grothmann and Patt (2005) seem to be amongst the first to point out that psychological factors were missing in the analysis of individual adaptive capacities. Until then, it was implicitly assumed that socio-economic and institutional factors are sufficient (Yohe and Tol 2002, Adger 2003, Smit and Pilifosova 2003, Smit and Wandel 2006). Since then, it has been widely acknowledged that psychological and cultural dimensions matter, but quantifying such aspects remains a challenge (Adger et al. 2009, Adger 2010, Adger et al. 2013).

Regarding the importance of self-efficacy for individual adaptive capacity, Grothmann and Patt (2005) discuss two case studies, one from Germany and one from Zimbabwe. However, the case study from Zimbabwe is more anecdotal so we focus here on the German case, where residents' proactive adaptation to the risk of flooding is investigated. Two models are compared: A socio-economic model and a socio-cognitive model. The former includes age, gender, human capital, net income, and whether the resident is a tenant or the owner of the house as explanatory variables. The latter includes only various perceptions, such as risk and self-efficacy. Interestingly, the latter model has more explanatory power than the former. In Ethiopia, Gebrehiwot and van der Veen (2015) investigate risk reduction measures of Ethiopian smallholder farmers. Also here, perceptions such as perceived vulnerability, severity of consequences, and especially self-efficacy, explain why some farmers undertake adaptation measures and others do not. A caveat of the survey based studies is the potential endogeneity of self-efficacy. It is not clear from the studies cited above, whether self-efficacy is purely a cognitive bias or whether it reflects some unobserved, objective factor of individual adaptive capacity. To take the Ethiopia study as an example, differences in objective farming skills could potentially explain why some farmers feel more vulnerable and why they report lower self-efficacy regarding the implementation of specific risk-reduction measures.

### 2.2. Experimental Evidence

In the psychological literature, the problems of reverse causality and omitted variable bias are commonly avoided by the use of experiments (Bandura 1997, 2012). This method has also gained popularity in economics and a recent study has investigated an effect of self-efficacy on the investment behavior of smallholder farmers in Ethiopia (Bernard et al. 2014). In this study, Bernard et al. (2014) raised the self-efficacy of a group of farmers by showing them a documentary about business success stories of social peers. The control group watched an uninformative program. Shortly after the treatment, the treatment group showed a significantly higher investment behavior than the control group, because observing successful peers had raised their aspirations. The valuable contribution of this experiment is the robust quantification of a causal effect of self-efficacy. However, it is *a* causal effect – not *the* causal effect. The authors are interested in the effect of higher aspirations, which they achieve by raising the farmers' self-efficacy. However, self-efficacy does not only affect aspirations, but as we discuss below, it affects also motivation and resilience to adversity, so the effect identified by Bernard et al. (2014) can be seen as a lower bound of the full effect. Unfortunately, self-efficacy needs very long to develop, so an experimental study on the full effect of self-efficacy seems prohibitively costly. In the context of climate change, or even weather shocks, an experimental treatment seems even more difficult.

### 2.3. Natural Experiments

A third option for research are natural experiments (Rosenzweig and Wolpin 2000). Natural experiments have the potential to overcome the constraints of either of the other approaches but they are hard to find. Wuepper and Sauer (2016) and Wuepper and Drosten (2016) argue that historic experiences can sometimes be used as instruments for current self-efficacy levels. In Ghana, they use colonial experiences with cocoa cooperatives to explain why some farmers inherited higher self-efficacy regarding contract farming from their parents; and they use historic farming systems to explain why some farmers inherited higher self-efficacy regarding agricultural investments. The big difficulty with their approach is the exclusion restriction: They argue that the colonial and pre-colonial experiences only affected the degree of inherited self-efficacies, but not the inheritance of other forms of capital, such as financial or social capital. However, even though they perform many tests to proof their assumptions, the exclusion restriction clearly is their studies' Achilles' heel.

### 2.4. Theory

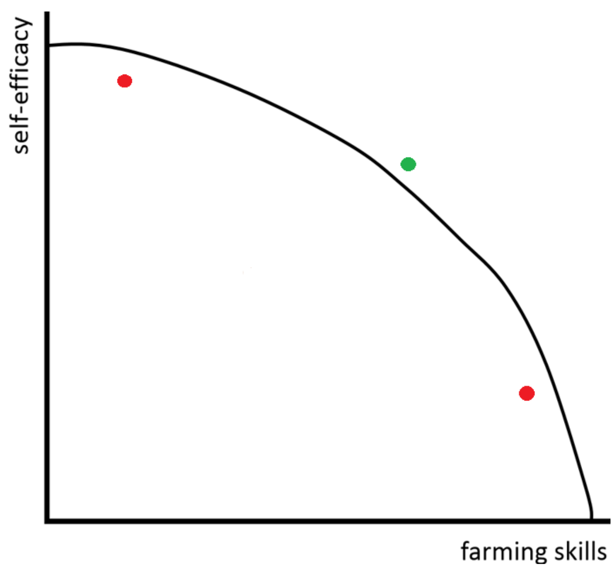
Let us briefly consider a bit of self-efficacy theory before we start with our analysis. Especially, we need to focus on one detail: What exactly is the relation between objective capabilities and subjective self-efficacy?

It is important to note that self-efficacy cannot compensate for a lack of ability (Pajares 1997, 2002). Self-efficacy determines what an individual can do given her abilities. It affects how effectively individuals can orchestrate their capabilities, and it affects the likelihood that skills are acquired (Bandura 1997, 2012). Unless decision-makers belief they can produce desired effects by their actions, they have little incentive to act. If they belief they can, this will influence the goals they pursue, how much effort they put forth, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-aiding or self-hindering, how they feel, and what they can accomplish (Bandura 1997). Thus, self-efficacy is usually the more severe constraint than objective ability (Pritchett and Kapoor 2009). An individual with low self-efficacy in a domain avoids that domain and thus never has the chance to increase her self-efficacy. If the individual is externally forced to engage in a domain with low self-efficacy, she is unlikely to be successful, because she is unlikely to invest sufficient effort and quite likely to give up early (Bandura 1997). In contrast, an individual sufficient self-efficacy is likely to acquire all necessary skills, given sufficient time and opportunity (Bandura 1997).

To clarify the short term relationship between objective skills and subjective self-efficacy beliefs, consider figure 1. The dots represent farmers and the curve represents the threshold at which the adaptive capacity of a farmers is sufficiently high to make the adoption of an innovation profitable.

Left of the curve, adoption is not profitable, right of the curve it is. The farmer all to the left does not have sufficient adaptive capacity, despite high levels of self-efficacy. She simply lacks the required skills (currently). The farmer all the way on the right also does not have sufficient adaptive capacity. In contrast to the first farmers, her constraint is the belief in her ability, even though she is the most skilled farmer in the figure. Only the farmer in the middle has sufficient adaptive capacity to profitably adopt the innovation. She neither has as much self-efficacy as the first farmer, nor is she as skilled as the second. However, she is sufficiently skilled and efficacious. With experience, both her skills and her self-efficacy will further improve as a function of her experiences.

**Figure 1. Self-Efficacy, Farming Skills, and Adaptive Capacity**




Notes: Adaptive Capacity increases from bottom left to top right. The frontier shows from which point on the adoption of an innovation is profitable. Only with sufficient objective skills and subjective beliefs does a farmer have sufficient adaptive capacity. Once she is right of the frontier, feedback effect move her further away from the frontier.

In the Annex, we provide sketch of the model of Just et al. (2009), which captures the basic mechanism how confidence in one's ability can affects profits and risks

## 3. Data

Before we can begin with our analysis, we briefly describe our sampling strategy (3.1) and how we capture the farmers' self-efficacy (3.2), their farming skills (3.3), and the rainfall variation they are exposed to (3.4). Then, we describe the technology which adoption we study (3.5) and show descriptive statistics (3.6).

### 3.1. Sampling

The data for this study comes from a survey conducted by the first author in 2013 in southern Ghana. Two sampling strategies were employed, to make the sample representative for all pineapple farmers in Ghana.

There exist reliable statistics on the export certified farmers in Ghana (Kleemann and Abdulai 2013, Wuepper et al. 2014). Thus, a three stage stratified sampling procedure was feasible, starting with the districts were most pineapples are produced (in the Eastern Region, the Central Region, and Greater Accra), followed by the farming groups that are certified to export pineapples, and finishing with a proportional sampling of individual farmer according to the number of pineapple producers. For non-certified pineapple farmers, there are no reliable statistics available, so the sampling is based on the information provided by development agencies and extension agents. When selecting non-certified farmers without lists, special emphasis was placed on the representativeness of the farmers, so as not to disproportionally sample "easier to reach" farmers.

The final sample size is 398 farmers, of whom roughly have been export certified at some point in time and the other half was never certified.

The data we use for the following analysis is pseudo-panel data and includes 5 periods (2009-2013) for the sampled 398 farmers. Obviously, this implies a certain degree of recall-error but this risk should not be overestimated as there is "true" panel data for about half the sample for two periods (periods 2009 and 2013 for the export certified farmers) which indicates that recalls are reliable. In fact, no farmer contradicts herself with clearly inconsistent answers in the two periods.

Of course, errors can still occur within the five periods, however, as long as this recall error is not systematically correlated with our variables of interest (especially with self-efficacy, e.g. farmers with higher self-efficacy being subject to less recall-error than farmers with lower self-efficacy) our results are unbiased. In general, measurement errors are likely to reduce the significance of all our estimates, so –if at all – we are less likely to find spurious effects with our data.

### 3.2. Self-Efficacy

Capturing a person's self-efficacy is a major empirical challenge (Bandura 2012, 2015). To do so Gebrehiwot and van der Veen (2015) ask Ethiopian farmers "How confident do you feel in general about your ability to protect yourself and your productive assets from severe drought risk" and "how confident do you feel you can perform the following risk reduction measures", on a scale from 1 = not confident to 7 = very confident. In Ghana, a pilot study suggested that asking about someone's self-efficacy directly could result in biased results. Instead, we chose to ask questions that were less obviously aimed at the farmers' self-efficacy and that could be used together as a factor variable in the analysis. The first question we asked was open ended about the two main income determinants in the last two years (roughly the growing time for pineapple). We then categorized the answers into categories, based on whether the named factors are under the control of the farmers or not. As a robustness check, we tried out different systems, i.e. using only the first answer (our favorite, because this shows what came to mind first), using an average of both answers, and using the higher of the two answers. Table 1 shows answer examples and how we coded them. Answers such as "I received training", "the productivity of my farm", or "my yields" were coded as "ambiguous", as they do not clearly reveal, whether e.g. the farmers meant "training was provided" (outside of control), or "I engaged in training and learned new skills" (under control), or how much she attributes yields and productivity to her own effort and skills. In contrast, most named determinants are clear cut, such as regular weeding (clearly a choice), or rainfall (clearly not a choice).

The second question aims at the relationship of the farmer with her environment: How much do you agree with the following statements: (A) "Nature provides what the farmer needs" or (B) "It is the task of the farmer to make nature productive". The possible answers were strongly agree with A, slightly agree with A, slightly agree with B, strongly agree with B, and other:___. Individuals with low self-efficacy in a domain prefer to think that they are provided for, either by another individual, a group, or in this case, their environment (Bandura 1997), which is a kind of "motivated belief" (Laajaj 2014, Bénabou 2015). This also showed up in the pilot study, in which some farmers insisted that their natural environment provides everything, so there is no need to use any inputs such as fertilizer.

**Table 1. Categorization of Mentioned Income Determinants**

| Under Control (3) | Ambiguous (2) | Outside Control (1) |
|---|---|---|
| agricultural practices | training | rainfall |
| seriousness | productivity | prices |
| regular weeding | yields | costs |
| learning | | diseases |

Notes: The table shows exemplary answers to the open question about the two main income determinants in the last two years. The answers were translates into an index according to whether the mentioned determinants were under the control of the farmer (1), outside of her control (3), or whether the answer was not clearly within either category (2).

The third question aimed at the planning horizon of the farmers. For a farmers who believes that she cannot achieve ambitious goals, there is not much incentive to plan for very long (Pritchett and Kapoor 2009, Banerjee and Duflo 2011). Of course, the feeling of agency that comes from self-efficacy does not only empower individuals – it can also be a weight on their shoulders (Fernandez et al. 2015). However, self-efficacy is a potential explanation for self-control problems, (a) because it reduces stress (Bernheim et al. 2013, Mani et al. 2013, Haushofer and Fehr 2014), and (b) because it makes goals more certain, lowering the discount rate (Mullainathan 2005, Datta and

Mullainathan 2014). The question was: “How far do you usually plan ahead when making major farming decisions?” The answers ranged from (a) one month to (f) longer than my lifetime.

Very likely, the categorized income determinant, and the reported nature relationship and planning horizon, all capture more farmer characteristics than only self-efficacy. The named income determinants do not only reflect the self-efficacy of a farmer but also her actual context; her planning horizon likely captures resource constraints too; and her nature relationship could also be correlated with the knowledge and skills of the farmers. To reduce the noise in these variables, we will use a factor variable, based only on the common variation in our three self-efficacy indicators. To furthermore reduce the risk of omitted variable bias, we sequentially add a large set of control variables in our specifications, which we describe in the next section. Directly below, we already explain how we construct one of these important control variables: The skills of the farmers. A first test for our self-efficacy factor variable is a simple regression on current and past determinants. Table 2 shows that only three variables explain the self-efficacy of the farmers. Consistent with the idea that self-efficacy is culturally inherited from one’s ancestors, we find that it is best explained with one’s ancestors having lived in an environment where cereals were crop of choice and where during colonial times, imposed cocoa cooperatives were more successful. As Wuepper and Drosten (2016) show, cereals reward agricultural investments more than roots, tubers, or trees, and thus, cereal farmers were incentivized to try out investments and to develop investment self-efficacy in response to positive experiences. Similarly, Wuepper and Sauer (2016) find that colonial cocoa cooperatives were likely to fail in regions where soils and rains where suboptimal, and where the distance to the colonial railroad increased transaction costs. Thus, in some regions, the farmers had success experiences with a modern value chain, that encouraged investments, whereas others experienced failure. This too shaped the self-efficacy of the descendants. Somewhat surprising is the finding that education lowers the self-efficacy of the farmers. One explanation is that education was increased by the activity of Christian missionary schools Wantchekon et al. (2015), which are found to have negatively affected social capital and self-efficacy in Ghana (Wuepper and Sauer 2016). In that case, the education estimate would be spurious. Alternatively, it is possible that indeed, more educated farmers have lower self-efficacy because they are more aware of difficulties or because the education system is truly demotivating, which however, cannot be inferred from our available data. Important to note: Having more farming skills and family labor only slightly correlate with our self-efficacy measure and external circumstances such as prices, city distance, rainfall, or contract farming are all insignificant predictors. Later in the analysis, we will use the historical roots of self-efficacy as a natural experiment to test for unobserved heterogeneity biasing our previous estimates.

**Table 2. Explaining Self-Efficacy**

| Dependent variable | Self-efficacy |
|---|---|
| hist. advantage cereals | 0.508*** |
| | (0.167) |
| hist. coops success | 0.419*** |
| | (0.136) |
| land | 0.0296 |
| | (0.0356) |
| education | -0.182*** |
| | (0.0319) |
| Farming skills | 0.0902* |
| | (0.0456) |
| age | 0.0336 |
| | (0.0417) |
| family labor | 0.0656* |
| | (0.0363) |
| contract farming | 0.0354 |
| | (0.0547) |
| export certified | 0.0150 |
| | (0.0677) |
| distance city | -0.218 |
| | (0.287) |
| local price | -0.0183 |
| | (0.0371) |
| company price | -0.00883 |
| | (0.0324) |
| rainfall | 0.0114 |
| | (0.0156) |
| soil | 0.0151 |
| | (0.0585) |
| district fe | yes |
| N | 1990 |
| R-sq | 0.35 |

Notes: Table shows estimated coefficients and standard errors (in brackets, clustered at districts). The model is OLS. Advantage cereals is the comparative historical suitability of an area to grow cereals, based on data from the FAO GAEZ database. Coop success is a measure of the regional success-rate of colonial cocoa cooperatives and missionary schools is the number of historic missionary schools around the farms, based on data from Cazzuffi and Moradi (2010), and Cogneau and Moradi (2014).

### 3.3. Farming Skills

An important question is whether self-efficacy is a bias or whether it actually reflects true ability. For Ghana's pineapple farmers, the best knowledge predictor is experience (Conley and Udry 2010). Formal education could also be a problematic variable if farmers with higher self-efficacy are more likely to be schooled, as found in Ethiopia (Bernard et al. 2014). To make our farming skill variable as comparable as possible to our self-efficacy variable, we base it on different measures of experience. First, we use how many years a farmer already produces pineapples. Secondly, we use self-reported expertise in comparison to her social peers. As our farming skills variable clearly correlates with the age of the farmers, we compare farmers who are older than average with farmers who are younger than average in table 3.

**Table 3. Differences between younger and older farmers**

| | younger | older | t-test |
|---|---|---|---|
| start pineapple | 2003 (6.171) | 2000 (7.776) | 9.790*** |
| start farming | 2002 (7.076) | 1997 (10.015) | 12.187*** |
| stated risk taking ability | 3.653 (.748) | 3.616 (.801) | 1.060 |
| perceived risk payoff | 3.766 (.600) | 3.805 (.597) | 1.480 |
| Planning horizon | 2.922 (1.816) | 3.349 (1.822) | 5.223*** |
| Off-farm income | .164 (.102)) | .147 (.101) | 3.700*** |
| experiment | .759 (.763) | .736 (.595) | 0.737 |
| experiment success | 3.713 (2.615) | 3.921 (2.600) | 1.763* |
| education | 2.756 (1.120) | 2.642 (1.269) | 2.124** |

Notes: The table compares farmers who are older than average with farmers who are younger than average. This comparison is important because our farming skills variable is highly correlated with experience, which in turn is correlated with age.

It can be seen that older farmers have unsurprisingly started farming earlier and also pineapple farming. They do not show generally different risk perceptions or attitudes but they plan longer, which is perhaps surprising, and they have less off-farm income, which is again consistent with expectations. Interesting for our study is the fact that older farmers are not more likely to experiment with new technology (which would suggest self-efficacy) but they are more likely to be successful (which suggest more skills).

### 3.4. Rainfall

Another crucial decision is what data to use to measure rainfall. There are only few weather stations in most Sub-Saharan African countries (Chaney et al. 2014), so in our research area in Ghana, there are only three. Model based data is available online (e.g. http://harvestchoice.org/products/data) but the effect of rainfall is even different within very small geographic units, depending on topography and soils. Finally, as suggested by Protection Motivation Theory, we need to pay attention to whether differences in risk perceptions affect our estimation of the self-efficacy effect (Floyd et al. 2000, Kroemker and Mosler 2002). By using farmer reported rainfall, we obtain a farm-specific measure for rainfall, of which we can be sure that the farmer perceives it exactly as we use it in the model. Because rainfall is a pivotal production input to grow pineapples, the farmers are well informed about it (Delavande et al. 2011). We show in the appendix that measured and reported rainfall are mostly similar, but that reported rainfall has much more variation. An interesting feature of our data is that it ranges from 2009 to 2013, a period starting with relatively high rainfall, which declines from year to year (starting at about 1073mm in 2009 and declining to 743 in 2013), causing many farmers to worry about future rainfall. In the annex, we provide a comparison of measured and reported rainfall.

### 3.5. The Technology

There are two technologies that could help the farmers to mitigate the adverse effects of low rainfall, of which only one is actually used. The obvious candidate would be irrigation, but this technology is mostly absent on the farms of smallholder farmers. A far less costly and complex technology is mulching. Mulching describes covering the bare soils to avoid direct evaporation of water, and to suppress weeds, which would compete with the pineapples for water and increase water loss through evapotranspiration (Erenstein 2003, Dzomeku et al. 2009, Snapp and Pound 2011, Kleemann and Abdulai 2012, Wuepper et al. 2014). The materials used for mulching differ widely, depending on material availability and affordability. The main materials used are either black plastic foil, which can be bought specifically for mulching, or organic materials, usually crop residues. A few farmers also find creative other ways and use materials such as old clothes to cover the ground of their fields. For this study, we do not restrict mulching to any particular material, as long as it is a non-living soil cover that is used to conserve soil moisture and prevent the growth of weeds. In the model, adoption is defined binary, as one if the farmer uses mulching on at least one plot, and zero otherwise.

### 3.6. Descriptive Statistics

In table 4 we present an overview over our main variables, separately reported for farmers with higher than average self-efficacy (HSE) and lower than average self-efficacy (LSE). It can be seen that farmers with higher self-efficacy are generally more likely to mulch, are more skilled in growing pineapples, are parts of networks where mulching is more diffused and live closer to a city. When we look at the correlation between farming skills and self-efficacy, we find only 12% correlation. This comes from the fact that self-efficacy has many sources, and objective capabilities are only a small part (Bandura 1997).

Table 5 presents a first comparison of farmers with different levels of self-efficacy regarding their perceptions, attitudes and general innovation behavior. It can be seen that all variables differ significantly between the farmers according to their self-efficacy. First of all, farmers with low self-efficacy perceive their external constraints as more severe, consistent with self-efficacy theory (Bandura 1997, 2012), except for the variable “market access”, which shows why self-efficacy is not the same as someone’s locus of control. Because the farmers with higher self-efficacy are more ambitious, they feel market access to be more difficult than the farmers with low self-efficacy, who do not even try to access more lucrative markets and are fine with selling locally at lower prices. Farmers with higher self-efficacy are also less risk averse, which has been found before (Krueger and Dickson 1994) and they are much more likely to have tried out an innovation in the past (almost double as likely) and also much more likely to have been successful with it. This too, is suggested by self-efficacy theory (Bandura 1997, 2012).

**Table 4. Main-Variables and Descriptions**

| Variable | Description | LSE | HSE |
|---|---|---|---|
| mulching | Whether the farmer mulches in a given period | .34<br>(.47) | .51<br>(.50) |
| income det. | Reported main income determinants in the last two years, coded according to nature (see table 2) and standardized | -.72<br>(.60) | .73<br>(.75) |
| nature attitude | How much the farmer beliefs that nature provides everything or that she needs to make nature productive (1 - 4) | 1.78<br>(1.04) | 2.54<br>(1.42) |
| planning | Self-reported planning horizon of the farmers (from only today to children's life, 1 - 6) | 2.71<br>(1.67) | 3.54<br>(1.88) |
| self-efficacy | Factor-variable from income det., nature attitude and the reported planning horizon | -.75<br>(.63) | .89<br>(.40) |
| drought | Opposite of reported rainfall (1 – 7), here: average | 2.11<br>(1.44) | 2.32<br>(1.35) |
| farming skills | Factor-variable from the years of pineapple production and self-reported expertise | -0.06<br>(0.09) | 0.06<br>(1.04) |
| farmland | Area of all fields (in hectares) | 10.08<br>(10.71) | 11.62<br>(12.52) |
| pineapple land | Area of pineapple fields (in hectares) | 3.24<br>(3.71) | 4.57<br>(6.33) |
| age | Age of the household head (years) | 43.17<br>(10.43) | 43.40<br>(11.20) |
| labor | Number of family members who work on farm | 1.50<br>(1.39) | 2.08<br>(2.04) |
| contract | Whether the farmer has a contract with a company (binary in the analysis, here %) | 17<br>(38) | 19<br>(39) |
| network | How much mulching was diffused in the social network of the farmer in the last period (%) | 37<br>(31) | 49<br>(33) |
| dist company | Distance between farm and next company (km) | 41<br>(33) | 42<br>(31) |
| dist city | Distance from farm to Cape Coast or Accra (km) | 40<br>(30) | 33<br>(15) |
| insurance | Whether or not the farmer has any insurance | .49<br>.50 | .36<br>.48 |
| inform-ins. | Whether the farmer is part of an informal insurance network or has friends who help in emergencies | .13<br>.34 | .12<br>.32 |

Notes: LSE denotes low and HSE denotes high self-efficacy. Statistics show mean and standard deviation.

**Table 5. Comparing Attitudes, Perceptions, and Behavior of the Farmers**

| Variable | Description | HSE | LSE | t-statistic |
|---|---|---|---|---|
| risk preference | Factor-variable from an experiment and various self-reports regarding risk | .148 (.580) | -.145 (.893) | 8.666*** |
| price constr. | How constraining are market prices? (1-6) | 2.340 (1.141) | 2.781 (1.061) | 8.927*** |
| credit constr | How difficult is credit access? (1-6) | 3.431 (1.067) | 3.567 (.833) | 3.164*** |
| labor constr. | How scarce is labor? (1-6) | 2.021 (1.053) | 2.645 (1.059) | 13.178*** |
| weather constr. | How problematic is the weather? (1-6) | 2.416 (1.171) | 2.308 (.954) | 2.252** |
| pests constr. | How high is risk of pests? (1-6) | 1.870 (.916) | 2.346 (1.174) | 10.067*** |
| insects constr. | How high is the risk of insects? (1-6) | 1.665 (.831) | 2.038 (.981) | 9.131*** |
| plants constr. | How scarce is planting material? (1-6) | 1.577 (.842) | 1.831 (.989) | 1.893*** |
| market constr. | Is market access a problem? (1-6) | 2.732 (1.397) | 2.554 (1.256) | 2.960*** |
| trials | Have you tried a new technology in the past? (1/0) | .963 (.783) | .538 (.498) | 14.463*** |
| Success of trials | How successful were these trials? (1 - 6) | 4.695 (2.073) | 2.945 (2.785) | 15.874*** |

Notes: table shows means and standard deviations (in brackets) as well as t statistics of whether the difference in means in significant between farmers with high (HSE) and low investment self-efficacy (LSE). Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

## 4. Analysis

Our analysis consists of two parts. First, we treat self-efficacy as an exogenous variable. Then we test this assumption using 2SLS. Our results suggest that our OLS estimates are unbiased and that self-efficacy is a robust determinant of the response to adverse weather.

### 4.1. Baseline

We begin our analysis with variations of the following model: The dependent variable is the probability that mulching is adopted ($P(Y_{ijt})$). This is explained with the self-efficacy of the farmers ($SE_{ijt}$), their farming skills ($S_{ijt}$), the drought severity in the last period ($D_{ijt-1}$), and a vector of individual control variables ($X_{ijt,t-1}$) and especially we compare the effects of two interaction terms: The interaction between having experiences a drought and the individual's self-efficacy $(D\ x\ SE)_{ijt}$ and the interaction between having experienced a drought and the individual's farming skills $(D\ x\ S)_{ijt}$. We always include district and year fixed effects $(\alpha_j + \delta_t)$, to control for unobserved district characteristics and time trends. We thus estimate variants of the following model:

$$P(Y_{ijt}) = \alpha_j + \delta_t + \beta_1(D\ x\ SE)_{ijt} + \beta_2(D\ x\ S)_{ijt} + \beta_3 SE_{ijt} + \beta_4 S_{ijt} + \beta_5 D_{ijt-1} + \beta_6 X_{ijt,t-1} + \varepsilon_{ijt} \quad (1)$$

Our First hypothesis is that farmers with higher self-efficacy and more farming skills respond significantly stronger to the experience of drought than farmers with lower self-efficacy and less farming skills. Our second hypothesis is that self-efficacy is more important than farming skills in affecting the drought response.

Table 7 shows 4 specifications. In the first, we only include self-efficacy, farming skills, and whether a farmer experiences a drought, and no interaction effects. We control for the credit access of the farmer, her farm size, age, education, labor, and where she sells her produce. It can be seen that both self-efficacy and farming skills positively affects the adoption of mulching, and notably, the experience of drought reduces the probability that mulching is adopted. In specification 2, we interact self-efficacy and farming skills with the experience of drought and find that farmers with higher self-efficacy are significantly more likely to respond to decreasing rainfall with the adoption of mulching than are farmers with more farming skills. In specification 3, we divide farmers into high and low self-efficacy, relative to the average. It can be seen that farmers with higher than average self-efficacy are likely to respond to the experience of drought with the adoption of mulching, whereas the sign is even negative (albeit insignificant) for farmers with lower than average self-efficacy. In specification 4, we also divide the farmers into having more or less than average farming skills. It can be seen that high self-efficacy continues to be the best predictor for why some farmers respond to decreasing rainfall with the adoption of mulching but it can also be seen that farmers with less farming skills are significantly less likely to adopt mulching after decreased rainfall than farmers with more farming skills.

**Table 7. Self-Efficacy, Farming Skills, and Climate Change Adaptation**

| spec. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | logit | logit | logit | logit |
| Dep.var. | adoption mulching | adoption mulching | adoption mulching | adoption mulching |
| self-efficacy (SE) | 1.759*** | 0.879 | | |
| | (0.386) | (0.572) | | |
| Farming skills (S) | 2.379*** | 1.869*** | | |
| | (0.399) | (0.584) | | |
| SE x drought | | 0.359* | | |
| | | (0.184) | | |
| S x drought | | 0.0806 | | |
| | | (0.201) | | |
| high SE x drought | | | 1.538** | 1.232** |
| | | | (0.620) | (0.574) |
| low SE x drought | | | -0.0731 | -0.170 |
| | | | (0.624) | (0.576) |
| high S x drought | | | | -0.137 |
| | | | | (0.215) |
| low S x drought | | | | -0.551** |
| | | | | (0.217) |
| drought | -0.839** | -0.309 | | |
| | (0.332) | (0.303) | | |
| controls | yes | yes | yes | yes |
| district fe | yes | yes | yes | yes |
| year fe | yes | yes | yes | yes |
| $pR^2$ | 0.41 | 0.42 | 0.41 | 0.40 |
| AIC | 936.9 | 920.8 | 931.9 | 940.7 |
| BIC | 1065.6 | 1060.7 | 1055.0 | 1075.0 |
| N | 1990 | 1990 | 1990 | 1990 |

Notes: Table shows estimated coefficients (marginal effects) and standard errors (in brackets). The model is probit. The observations come from 398 farmers in 5 periods. Standard errors are clustered at the farmer level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***). Throughout we control for credit access, farm size, education, age, labor, and marketing channel.

### 4.2. Robustness Analysis

To test the robustness of our estimates, we perform 2SLS regressions, instrumenting the self-efficacy of the farmers with certain historical experiences of their ancestors. We follow the idea of Wuepper and Drosten (2016) who argue that a long-term dependency on cereals gradually built up investment self-efficacy because cereals have higher returns on investment than roots, tubers, or trees. Thus, they argue, a historic dependency on cereals is a feasible instrument for the self-efficacy of current non-cereal farmers, such as pineapple farmers. They show that indeed, farmers whose ancestors where cereal farmers have higher investment self-efficacy today. However, demonstrating that the exclusion restriction holds is a bit tricky. If one's ancestors were growing cereals instead of e.g. roots, and if one's ancestors had higher investment self-efficacy, these factors could have led to an accumulation of productive assets from which the current farmers benefit, over and above of their inherited self-efficacy. In order for the approach of Wuepper and Drosten (2016) to work, Malthusian dynamics must have led more successful farmers to have larger families but individual ancestors cannot be better off. We test this by regressing five outcomes that could indicate differences in inherited assets. The first is inherited land in hectares. Farm-size is likely to be partially an outcome of the self-efficacy and performance of a farmer but inherited land under traditional tenure rights is a feasible indicator for whether cereal farmers have given more assets to their children. Specification 1 in table 8 suggests that they did not. Specification 2 shows that descendants of cereal farmers also do not differ from the descendants of other farmers by their educational attainment and specification 3 shows that they are neither more likely to have a bank account. All this support the historical dependency on cereal farming as a feasible instrument for the self-efficacy of the pineapple farmers. Another test is to regress the local pineapple price on the regional dependency on cereal farming. Richer regions should have a higher pineapple price, but specification 4 shows that the price is not correlated with historical cereal farming. Thus, we cannot reject the hypothesis that a historic dependency on cereal farming only affects the current generation of pineapple farmers through their self-efficacy.

**Table 8. Instrument Falsification Tests**

| spec. | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| model | OLS | OLS | OLS | OLS |
| dep. var | land endowment | Education level | bank account | pineapple price |
| Historically Cereal Farming | 0.000418 (0.00550) | -0.0754 (0.0487) | 0.00410 (0.0183) | -0.0425 (0.0344) |
| district FE | yes | yes | yes | yes |
| R-sq | 0.109 | 0.047 | 0.082 | 0.197 |
| N | 398 | 398 | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). The model is probit. The observations come from 398 farmers in 5 periods. Standard errors are clustered at the farmer level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

To avoid the danger that historically, societies with higher self-efficacy sorted into cereal farming, we use as an instrument not whether one's ancestors actually participated in cereal farming, but whether her biogeographic context created an incentive to do so. To this end, we use data from FAO's GAEZ database and use GIS software to compute where cereals had a comparative advantage in comparison to growing roots or tubers. The specific crops that we use are maize as a cereal and yams as a roots or tuber. Furthermore, we include another instrument, which we borrow from Wuepper and Sauer (2016): Rainfall in the 1930s. The 1930s were a crucial time period for the farmers in Southern Ghana, because the colonial government wanted to promote the export of high quality cocoa to Europe and established cocoa cooperatives. The goal was to intensify cocoa production by means of this organizational innovation. Due to biogeographic factors such as rainfall, some of these cooperatives were successful, whereas other failed. Wuepper and Sauer (2016) find that this experience also shaped the self-efficacy of the pineapple farmers, and again, did not have any other long-term effect on the current generation of pineapple farmers. Thus, we use the regional advantage to grow cereals, or roots and tubers, and the local rainfall in the 1930s as instruments for the self-efficacy of the pineapple farmers. We estimate variants of the following model:

$$P(Y_{ijt}) = \alpha_j + \delta_t + \beta_1 \widehat{(D\ x\ SE)}_{ijt} + \beta_2 (D\ x\ S)_{ijt} + \beta_3 SE_{ijt} + \beta_4 S_{ijt} + \beta_5 D_{ijt-1} + \beta_6 X_{ijt,t-1} + \varepsilon_{ijt}$$
$$(D\ x\ SE)_{ijt} = \alpha_j + \delta_t + \beta_1 I_j + \beta_2 (D\ x\ S)_{ijt} + \beta_4 SE_{ijt} + \beta_5 S_{ijt} + \beta_6 D_{ijt-1} + \beta_7 X_{ijt,t-1} + \varepsilon_{ijt} \quad (3)$$

Where $P(Y_{ijt})$ is the probability that a farmer adopts mulching, $(D\ x\ SE)_{ijt}$ is an interaction term between self-efficacy and having experienced a drought in the last period, $(D\ x\ S)_{ijt}$ is an interaction term between the farmer's skills and having experienced a drought in the past period and $I_j$ is the vector of instruments, namely the historic, comparative advantage in growing cereals or roots and tubers, and the local rainfall in the 1930s.

**Table 9. 2SLS:**

**The Probability to Adopt Mulching after a Drought as a Function of Self-Efficacy**

| spec. | (1) | (2) | (3) |
|---|---|---|---|
| model | 2SLS | 2SLS | 2SLS |
| dep.var | adoption mulching | adoption mulching | adoption mulching |
| SE x D | 0.503*** | 0.502*** | 0.318** |
| | (0.129) | (0.144) | (0.141) |
| controls | A | B | A |
| effects | random | random | fixed |
| F excluded SE | 106.69 | 143.03 | 106.69 |
| F excluded SExD | 65.76 | 86.57 | 65.76 |
| R-sq | 0.03 | 0.06 | 0.03 |
| N | 1990 | 1990 | 1990 |

Notes: Table shows estimated coefficients and standard errors (in brackets). The model is probit. The observations come from 398 farmers in 5 periods. Standard errors are clustered at the farmer level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

Table 9 reports the first set of results. Controls A include age, education, self-efficacy, farming skills, and drought in the last period and controls B also include soil fertility, farm-size and distance to urban markets. The F test for the included instruments is always high, both for self-efficacy (F excluded SE) and for the interaction between self-efficacy and drought (F excluded SExD), so we do not need to worry about weak instruments bias. In the random effects specifications 1 and 2, the estimated response to drought of farmers with higher self-efficacy is slightly stronger than in the baseline model (0.5 in contrast to 0.36). In the fixed effects specification 3, the effect is almost the same as before (0.3 in contrast to 0.36). This suggests that our OLS estimates are not biased by unobserved heterogeneity but that it is important to control for time-invariant farmer characteristics.

**Table 10. 2SLS:**

**Comparing the Response as a Function of Self-Efficacy and Farming Skills**

| spec. | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| model | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| dep.var | adoption mulching | adoption mulching | adoption mulching | adoption mulching | adoption mulching | adoption mulching |
| SE x D | | 0.543*** | | 0.530*** | | 0.345** |
| | | (0.145) | | (0.160) | | (0.163) |
| S x D | 0.0141** | -0.0856*** | 0.0149*** | -0.0832*** | 0.00937* | -0.0541* |
| | (0.00558) | (0.0295) | (0.00561) | (0.0318) | (0.00564) | (0.0315) |
| controls | B | B | B | B | B | B |
| effects | random | random | random | random | fixed | fixed |
| F excl. SE | 107.30 | 107.30 | 144.87 | 144.88 | 107.30 | 107.30 |
| F excl. SExD | | 66.47 | | 90.58 | | 66.47 |
| R-sq | 0.01 | 0.03 | 0.01 | 0.03 | 0.01 | 0.03 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: Table shows estimated coefficients and standard errors (in brackets). The model is probit. The observations come from 398 farmers in 5 periods. Standard errors are clustered at the farmer level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

To return to our question whether farming skills or self-efficacy are more important for the adaptive capacity of the farmers, we include interaction terms of drought and farming skills in specifications 1, 3, and 5, and include both the interaction between farming skills and drought and self-efficacy and drought in specifications 2, 4, and 6. The interesting pattern that emerges is that as long as we do not control for the self-efficacy of the farmers, farmers with more farming skills are estimated to be more likely to respond to a drought with the adoption of mulching than their less skilled peers. However, if we control for the self-efficacy of the farmers, the estimated effect of the farming skills becomes significantly negative. This does not mean that farming skills really lower the probability of technology adoption but it suggests that it is mostly the self-efficacy of the farmers that determines their response to adverse weather.

## 5. The Income Effect of Self-Efficacy and Mulching

Our results above lead to two additional questions: First, what happens if farmers with low self-efficacy adopt mulching? Do they invest sufficiently and benefit from the adoption, or does it actually take high self-efficacy to benefit? Secondly, if we find that farmers with high self-efficacy are more than twice as likely to respond to adverse weather with the adoption of mulching, what is income effect of this difference?

**Table 11. The Income Effect of Self-Efficacy and Mulching**

| spec. | (1) | (2) | (3) |
|---|---|---|---|
| model | OLS | OLS | OLS |
| dep.var. | income | income | income |
| high SE x mulch | 0.108** | | 0.105** |
| | (0.046) | | (0.046) |
| low SE x mulch | 0.001 | | 0.002 |
| | (0.039) | | (0.039) |
| high SE x drought | | -0.104* | -0.119** |
| | | (0.058) | (0.059) |
| low SE x drought | | -0.170*** | -0.159** |
| | | (0.062) | (0.062) |
| controls | yes | yes | yes |
| district fe | yes | yes | yes |
| year fe | yes | yes | yes |
| R-sq | 0.19 | 0.19 | 0.19 |
| N | 1990 | 1990 | 1990 |

Notes: Table shows estimated coefficients and standard errors (in brackets). The model is OLS. The observations come from 398 farmers in 5 periods. Standard errors are clustered at the farmer level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***). Throughout we control for credit access, farm size, education, age, labor, and marketing channel. Prices are deflated to 2002 Ghana Cedis according to the consumer price index of the Ghana Statistical Office.

Table 11 shows 3 specifications. In the first, it is suggested that indeed, only farmers with high self-efficacy benefit from the adoption of mulching. The second specification suggests that farmers with low self-efficacy loose indeed significantly more income from adverse weather conditions than farmers with high self-efficacy. However, it can also be seen that all farmers loose and no group can fully mitigate the adverse effect.

In specification 3, we include the arguments together and find that controlling for different technology adoption responses of the farmers, the difference in income-reduction between farmers with high and low self-efficacy narrows but does not disappear. This suggests that it is not only the adoption of mulching that helps high self-efficacy farmers to reduce the adverse income effect of lower rainfall but they also do other things that are beneficial -which could be more regular weeding (reducing competition for soil moisture), using agroecological practices that support the plants and improve the soil (e.g. planting certain leguminoses), or improved soil preparation.

## 6. Discussion and Conclusion

Understanding individual adaptive capacity and technology adoption decisions are crucial information for climate policy (Tol et al. 1998, Di Falco et al. 2011, Dinar et al. 2012, Zilberman et al. 2012, Di Falco 2014, Barros et al. 2015). However, we currently do not fully understand technology adoption decisions (Foster and Rosenzweig 2010, Zilberman et al. 2012) and we know less about adaptive capacities (Grothmann and Patt 2005, Adger et al. 2009). Whereas a common assumption is that income is the main determinant of climate change impacts (Yohe and Tol 2002, Dell et al. 2012), there is some research suggesting that independent from resources and economic incentives and constraints, there are also psychological, social, and cultural factors that are important (Grothmann and Patt 2005, Adger et al. 2009, Jones and Boyd 2011, Gebrehiwot and van der Veen 2015). Above, we compare the effect of farming skills with the effect of self-efficacy and find that self-efficacy is the more important factor. We estimate that a farmer with high self-efficacy is more than twice as likely to adopt mulching in response to drought than a farmer with low self-efficacy. The estimated effect of self-efficacy is robust to using different models and specifications and survives 2SLS estimation. The implications for modelling climate change impacts (e.g. in Integrated Assessment Models such as discussed by Pindyck (2013)) are that adaptive capacity is likely to be overestimated in many regions of the world, because there are more constraints to the adoption of innovation than usually assumed in the models. Hertel and Lobell (2014) make a similar argument for economic constraints.

Barely recognized is the fact that self-efficacy could easily be incorporated into standard economic models. Similar to civic capital, self-efficacy is a kind of capital. It can be thought of as a Bayesian prior about one's abilities. However, in contrast to the standard Bayesian model, it is pseudo-Bayesian, in the sense that the prior biases how the posterior is updated. We can augment the model of Just (2002) to incorporate the causal effect of initial self-efficacy $p_t(x)$ on the posterior self-efficacy $P_{t+1}$:

$$P_{t+1} = \frac{p_t(x)^{R(l,p,z)} l(\theta | x(\, p_t(x)))^{L(l,p,z)}}{\int_{-\infty}^{\infty} p_t(x)^{R(l,p,z)} l(\theta | x(\, p_t(x)))^{L(l,p,z)} dx}, \qquad (3)$$

where $l$ denotes new information, $R$ and $L$ are the weights a person gives to her initial belief and new information, respectively, and $z$ reflects the costs of learning. It can be seen that initial self-efficacy impacts the development of subsequent self-efficacy but the mapping is not one to one. The research of Krishnan and Krutikova (2013), Bernard et al. (2014) and that discussed by Bandura (1997) shows how self-efficacy can be improved at different stages of personal development and in different contexts. Thus, one the one hand, self-efficacy is quite persistent throughout time, because individuals with low self-efficacy are unlikely to make the kind of experiences that would improve their self-efficacy. On the other hand, with external support, not only their own self-efficacy can be improved, but through social learning, the self-efficacy of their social peers can be improved to. In Ethiopia, Bernard et al. (2014) showed farmers a documentary about economically successful social peers, after which the farmers raised their aspirations and increased their investments. How strong and general this self-efficacy develops now depends on their experiences, which can of course be affected by direct support and policies. In India, Krishnan and Krutikova (2013) greatly improved the self-efficacy of school children, using a multidimensional program that provided experiences, mentoring, and contemplation. In an athletic context, Weinberg et al. (1981) nicely demonstrate how self-efficacy can be manipulated. Before an athletic competition, they told one group of subjects their competitors were varsity track athletes and they told the other group their competitors had a knee injury. The subsequent competition was rigged, so that all test subject lost. Consistent with the theory, individuals with high self-efficacy (who thought they were competing against an injured competitor) tried much harder to win than individuals with low self-efficacy (who thought they competed against a superior competitor). After losing, the individuals with high self-efficacy further increased their efforts whereas individuals with low self-efficacy further reduced it. The implication for communication with farmers is that criticizing current practices and increasing the worry about climate change and weather impacts can have a positive effect if the main constraint to behavioral change is a lack of problem awareness, but it can have a negative effect if the main constraint is in fact self-efficacy. Our research and that of others suggests that self-efficacy is an important behavioral determinant and thus, we recommend self-efficacy to be explicitly considered when one wants to achieve a change in agricultural practices.

We find suggestive evidence that farmers with higher self-efficacy also show other behavioral differences that allow them to reduce the adverse income effect of missing rainfall.

As a final note, self-efficacy and actual ability are rather complements than substitutes. This is important to keep in mind, because attempting to increase self-efficacy without empowerment is likely to lead to disappointment and a reduction in welfare. However, the combination of improving self-efficacy and the objective context is likely to have a much larger effect than a pure focus on technical skills, inputs, and infrastructure.

## Annex

A Theoretical Model

A self-efficacy model can look like the confidence-model of Just et al. (2009). Let us assume an innovation gives the outcome distribution $f(\pi|\mu,\sigma^2)$, where $\pi$ is the return on investment, and $\mu$ and $\sigma^2$ are its mean and variance. Thus, the investment risk is described by the variance of the expected profit.

Farmers with low self-efficacy belief the distribution is $f(\pi_i|\mu,\sigma_s^2)$ with $\sigma_s^2 > \sigma^2$ and farmers with high investment self-efficacy belief it is $f(\pi_i|\mu,\sigma_s^2)$ with $\sigma_s^2 < \sigma^2$. To see how these beliefs affect the actual profitability of the innovation, we start by assuming that utility is a simple function of profit, that it is concave, continuous, and differentiable, and that it can be approximated by the following Taylor series:

$$u(\pi(x,p)) = u(\pi(x,\mu)) + u'(\pi(x,\mu)) \cdot x \cdot (p-\mu) + {}^1\!/_2\, u''(\pi(x,\mu)) \cdot x^2 \cdot (p-\mu)^2 \qquad (1)$$

where $x$ is output, $p$ is price, and profits are $\pi(x,p) = px - C(x) - B$, where $C(x)$ are variable costs (with $C(0) = 0, C'(x) > 0$) and $B$ are the fixed costs. The farmers maximize their expected utility:

$$\max_x E\left[u(\pi(x,\mu)) + u'(\pi(x,\mu)) \cdot x \cdot (p-\mu) + {}^1\!/_2\, u''(\pi(x,\mu)) \cdot x^2 \cdot (p-\mu)^2 | \mu_s, \sigma_s^2\right] =$$

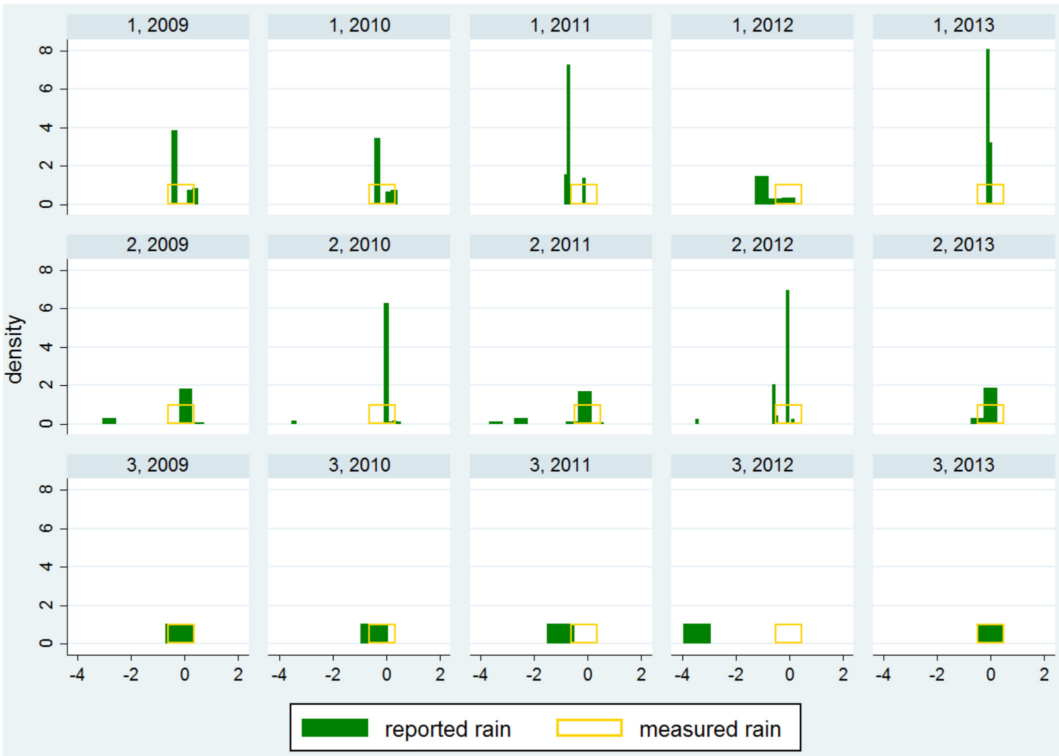
$$u(\pi(x,\mu)) + u'(\pi(x,\mu)) \cdot x \cdot (\mu_s - \mu) + {}^1\!/_2\, u''(\pi(x,\mu)) \cdot x^2 \cdot \sigma_s^2 \qquad (2)$$

After a few mathematical operations (and some qualifying assumptions), the following comparative static result can be obtained:

$$\frac{dx}{d\sigma_s^2} = \frac{-\left[{}^1\!/_2 \cdot ({}^{u''}\!/_{u'}) \cdot \left(\mu - C''(x)\right) \cdot x^2 - (-{}^{u'''}\!/_{u'})x\right]}{SOC} < 0 \qquad (3)$$

Where $x$ is still the agricultural output and $\sigma_s^2$ is still the believed risk of the innovation, and $({}^{u''}\!/_{u'})$ and $(-{}^{u'''}\!/_{u'})$ respectively denote measures of absolute risk aversion and prudence. Thus, perceived self-efficacy affects the actual risk and profitability of technology adoption.

### A Comparison of data from Ghana's weather stations and farmer reported rainfall




Notes: The comparisons above show standardized rainfall in the Eastern Region (1), the Central Region (2), and Greater Accra (3) between 2009 and 2013. Reported rain (green) is farmer specific whereas measured rain (yellow) is only per region. It can be seen that the general pattern is similar but more so in the years 2009, 2010, and 2013, and less so in the years 2011 and 2012.

# Chapter 4

# Moving Forward in Rural Ghana: Investing in Social and Human Capital Mitigates Historical Constraints

with Johannes Sauer[5]

**Abstract** It is often found that historic events have a persistent impact on economic outcomes. We surveyed 400 pineapple farmers in Ghana and find that indeed both the historic dependency on different crops and the experience of the trans-Atlantic slave trade still explain income differences in 2013. Based on this finding, we ask what characterizes the farmers who achieve higher than predicted incomes.

We find that such farmers are mostly enabled by social and human capital. A mediation analysis shows that about 30% of the income effect of social capital is explained by its effect on farming practices, especially the choice of growing a more demanding pineapple variety and using more inputs. In contrast, less than 10% of the education effect is mediated through this channel.

This chapter is currently under peer review with Economic History of Developing Regions.

**Contributions** David Wuepper performed the analysis and wrote the article. Johannes Sauer provided valuable comments and suggestions. The research question why history is not equally persistent for everybody was developed during a research discussion with Davide Cantoni.

---

Technical University Munich, Agricultural Production and Resource Economics

# 1. Introduction

In recent years, much has been learned about the historic roots of economic development (Nunn 2009, 2012, Alesina and Giuliano 2013, Nunn 2013, Spolaore and Wacziarg 2013).

In Ghana, Cogneau and Moradi (2014) find that different colonial policies of the French and the British in Togoland (now a part of modern Ghana) led to notably different literacy rates - which still persist today. Jedwab and Moradi (2015) investigate the effect of colonial railroad construction and find that even after the railroads lost their function (because they were not maintained and roads became the dominant transportation infrastructure) their economic effects persists.

Also in Ghana, Wuepper and Drosten (2016) develop an agent based model which suggests that the historic return on agricultural investment persists to shape expectations amongst farmers, which explain why some farmers invest more than others and thus have higher incomes. This is supported by empirical evidence from smallholder pineapple farmers. Wuepper and Sauer (2016) investigate what determines the performance of pineapple farming contracts between export-companies and smallholder pineapple farmers. They find that farmers facing the same economic incentives and constraints differ in their reliability because they have different beliefs about the payoffs from such a marketing channel and different amounts of social capital. Both variables are historically rooted: The farmers' beliefs come from historic experiences of their ancestors with colonial cocoa cooperatives and social capital has been persistently lowered by the activity of Christian missionary schools (see also Cazzuffi and Moradi (2012) for more information on the cooperatives and Claridge (1915) and Ward (1966) for assessments of the missionary impact in the Gold Coast).

So far, the focus in the literature was mostly on establishing the link between history and current economic outcomes. This was important, because history was not a standard part of development economics, especially not before the seminal work of Acemoglu and Robinson (2001) and

Acemoglu et al. (2002). Their research inspired much subsequent investigations (Austin 2008), such as whether history matters more because of institutional or cultural effects (Nunn and Wantchekon 2011, Alesina and Giuliano 2013). However, very little research has been conducted so far investigating when history does not persist and when it does not, why? The wealth of empirical evidence on the persistence of history can give the interpretation that economic outcomes are fully determined by the past and there is not much that individuals and policy makers can do. This would, however, be a stark misinterpretation, as it was simply necessary to first establish that history is persistent, in order to make the next step and analyze when it is not (Nunn 2013). In Africa, research about factors that mitigate historically inherited constraints is only represented by the explorative investigation of Nunn (2013) and the "side-results" of Nunn and Wantchekon (2011). Worldwide, it seems only Grosjean (2014) can be added for the US and Voigtländer and Voth (2012) for Germany.

Nunn and Wantchekon (2011) explain interpersonal mistrust in Africa with a long-term effect of the large slave trades, which persists to lower economic development until today. However, they also find that education and other individual specific variables explain more variation in trust, which suggests that fostering such other factors would help to overcome historically inherited mistrust. Nunn (2013) presents suggestive evidence that the adverse effect of the slave trades is weaker in countries with better domestic institutions, which suggests that institutional improvements could compensate for historical culture shocks. A similar finding is presented by Grosjean (2014) for the US, where she identifies historically determined cultural differences as explanation for different rates of violent crimes. Notably, she finds this persistence only in the US South, where property rights institutions and the rule of law were historically weaker than in the US North. In Germany, Voigtländer and Voth (2012) first establish a strong persistence of anti-Semitic values and beliefs between the 14th and the 20$^{th}$ century and then continue by demonstrating the effect to be much weaker in Hanseatic cities, which were involved in lucrative long-distance trade, in cities with faster population growth, and cities that were more industrialized in 1933.

This article contributes to the sparse empirical evidence on factors that allow better economic outcomes than historically predicted. It is also the first to focus on smallholder farmers in Sub-Saharan Africa and to consider human and social capital. Given the apparent importance of historical persistence and smallholder farming in developing countries around the world, the policy implications of this research could be important.

Because we only consider two (far away) points in time - the pre-colonial period and the presence – we are carefully considering possibly omitted variables and alternative channels of causation. For example, we focus on historic variables that shaped culture and thus also test whether colonial investments into infrastructure or education, or current tenure security matter, as they would point towards an institutional channel as well.

We focus on pineapple farmers in Ghana and test whether human and social capital, as well as exporting and more intense production explain why some farmers do better than predicted from the experiences of their ancestors.

First, we estimate the effect of two historical variables:

(a) The impact of the trans-Atlantic slave trade, which is established to have lowered economic development by reducing inter-personal trust, and

(b) A historical reliance on cereals as main crop, instead of hunting and gathering, fishing, tree-crops, roots, or tubers. This is found to have changed the expectation of farmers that agricultural investments are profitable, because of the historical profitability of such investments. Investigating the incomes of Ghana's pineapple farmers, investments are the most important income determinant. Thus, investment-increasing beliefs significantly increase incomes.

Then we use the historic variables to predict the current income of the farmers and compare the prediction with their actual income and investigate which farmers positively deviate from their historically predicted income.

We find that farmers who achieve higher incomes than predicted export their pineapples and produce more intensively. However, these characteristics are not the fundamental cause of their success in beating history. Instead, more intensive production is part of the mechanism and the correlation between beating the historical prediction and exporting is purely based on selection.

The fundamental causes that allow the farmers to achieve higher incomes are social and human capital.

As a short summary of the history in rural, southern Ghana, we might start with the accounts of George Peter Murdock (1959) and also consider Ward (1966), Macdonald (1898), Claridge (1915), and Austin (2005).

It was about the 13$^{th}$ century when people started to settle in the area of modern Ghana (Ward 1966). First, the Akan-speaking people (Ashanti, Akyem, Fanti) arrived from the north, then, at the start of the 17$^{th}$ century, Ewe, Ga and Adangme arrived from the east. Early trade established between Islamic traders from North Africa (who brought salt) and the states of Ghana's forest zone (that produced gold). The forest zone occupies the middle of the country and is bordered by the coastal zone in the south (where Fanti and Ga lived). The Savanna zone in the north lies outside our sample area. In addition to rich gold reservoirs, the forest zone is also more humid and has better soils than coast and Savanna. Perhaps for this reason, the largest society in this area, the Ashanti, developed the most powerful kingdom with more complex structures than its neighbors. Especially between the 17$^{th}$ and the 19$^{th}$ century, the Ashanti dominated the other societies and frequently invaded their territories to capture slaves.

The main subsistence source was determined by biogeographic factors. In the eastern region, farming systems where especially cereal based, whereas they where based on roots and tubers along the coast of the central region, were they where often complemented with fisheries.

The first Europeans to arrive where the Portuguese in 1471. They established a lucrative trade at the coast in which they exchanged European goods especially for gold (but also ivory, pepper and other valuable resources). They also brought exotic plants from their other colonies, which included oranges, lemons, limes, rice, sugar cane (from Asia) as well as maize, tobacco, pineapple, cassava and guava (from the Americas). Their trade-monopoly, however, was soon challenged by other European powers, beginning with the French in 1500. In 1553 the British entered the trade and later Danes, Germans, Swedes and Dutch would follow. For its most precious resource, the area was now called the "Gold Coast".

Even though gold continued to be a valuable commodity, the trade quickly developed towards focusing on an even more valuable resource: Slaves. Slavery was common in most of the Gold Coast's societies and slaves had been traded over the trans-Saharan trade route before (Claridge 1915). However, there was both a quantitative and a qualitative difference to the trans-Atlantic slave trade: First, coming from the colonies in Latin-America and the Caribbean there was an unprecedented surge in demand for slave-labor. Second, a vicious feedback-effect was started with this growing demand, because the incentive to capture and sell slaves disintegrated prior well-functioning societies (European traders never went inland to capture slaves themselves but they bought them from local traders). As analyzed by Lovejoy (2011), it was not only the financial support for aggressive slave-raiders that destabilized the societies in the Gold Coast. Societies and villages now had to protect themselves from capture, so they needed weapons, which they could get from the Europeans in exchange for slaves.

Nevertheless, slave-traders from Fante, Ga and Ashanti made great profits and especially the kingdom of Ashanti grew more powerful from year to year. With its growing power, its aspirations also grew and invasions into surrounding societies became an integral part of its expansionist policy (Ward 1966).

A highly visible European influence in the Gold Coast where the Christian missionaries. Already the first Portuguese explorers brought catholic missionaries. They were later followed by the protestant Basel mission, which founded a station at Aburi in 1847, the protestant Bremen mission, which worked east of the Volta and a Methodist mission, which started in the Fante area (Macdonald 1898, Ward 1966). A Wesleyan and a Scottish mission were established soon after (Cogneau and Moradi 2011).

To attract people to Christianity, it was soon discovered that the provision of formal education was most effective (Nunn 2010). Thus, Christian schools spread over the country, where scholars learned how to read, write and calculate.

In the 19th century, political development in Europe caused the British to change their politics in the Gold Coast entirely. They abolished the slave trade and turned to the so-called "legitimate trade" in agricultural commodities (especially cocoa but also maize, rice and palm oil) and natural resources (especially timber). This was obviously against the commercial interests of the Ashanti kingdom and the British, who by now completely dominated their European competition, decided to colonize the Gold Coast completely, to have proper control over the territory. After three heavy combats with the Ashanti, in 1867 Ashanti become officially British protectorate and the Gold Coast a British colony (Claridge 1915, Ward 1966).

The main export good of the Gold Coast was now cocoa. In the 20th century, due to strong demand from industrializing Europe, the cocoa trade accounted for 60 to 80% of the Gold Coast's exports (Austin 2005, Cazzuffi and Moradi 2010). To improve the quality of the produce, the colonial government decided by the 1920s to organize the farmers in cooperatives. This enabled them to provide credit, training and quality control (similar to modern contract farming). Much later, in the 1990s, the farmers of Southern Ghana began to cultivate pineapples for export (Conley and

Udry 2010). At first, this was a great business opportunity but in the 2000s, increasing competition, especially from Costa Rica, revealed the substantial risk of this value chain and when the European Union, the destination for Ghana's export pineapples, demanded a new variety, many companies and farmers went out of business. The surviving companies and farmers are still in a process of adaptation and especially to develop more modern and reliable farming and business structures. They have help from various development organizations and the government of Ghana (Wuepper 2014, Wuepper et al. 2014).

This article is structured as follows: In the next section (2), data and methodology are outlined. In section 3, the current incomes of Ghana's smallholder farmers are predicted with historical variables. In section 4, it is analyzed which farmers achieve higher than predicted incomes and why. In section 5 a few robustness checks are performed and in section 6, the results are discussed and the study is concluded.

## 2. Data and Methodology

The data comes from a survey conducted by the first author in 2013 in southern Ghana and the sources listed in table 1. For the survey, two sampling strategies where employed - to make the sample representative for all pineapple farmers in Ghana. There exist reliable statistics on the export certified farmers in Ghana (Kleemann and Abdulai 2013, Wuepper et al. 2014). Thus, a three stage stratified sampling procedure was feasible for those farmers, who export at least some of their pineapples. The first stage are the districts were most pineapples are produced (in the Eastern Region, the Central Region, and Greater Accra), the second stage are the farming groups that are certified to export pineapples, and the third stage is comprised of proportional sampling of individual farmer according to the number of local pineapple producers. For non-certified pineapple farmers, there are no reliable statistics available, so the sampling is based on the information provided by development agencies and extension agents.

The final sample size is 398 farmers, for whom we show descriptive statistics in table 2. It can be seen that these farmers are neither rich nor poor (average annual income in from pineapple was GHC 2171 in 2013, which was roughly US$ 1000). On average, they are about 44 years old, have finished junior secondary school, and have about 13 years of experience with growing pineapples. A third of them has access to credit, and one fifth of them has access to off-farm income. Half of the farmers uses chemical fertilizer and mulching (a technique to cover the soil to mitigate weeds and conserve soil moisture) and a few of them uses crop residues as organic fertilizer or material for mulching. Furthermore, about a third of the farmers uses the modern MD2 variety and also a third exports pineapples to the European Union (through delivering the pineapples to an exporting company).

**Table 1. Data Sources**

| variables | source |
|---:|---|
| farmer and farm characteristics | Own Survey |
| distances, topography | Own computations in GIS software, data from ArcGis and DivaGIS |
| soil suitability to grow different crops | Own computations in GIS software, data from FAO's GAEZ database |
| data on external slave trades in Africa | Nunn and Wantchekon (2011) |
| data on historic school locations (missionary and government) | Cogneau and Moradi (2011), Nunn (2010) |
| Colonial railroad tracks | Jedwab and Moradi (2012) |

## Table 2. Descriptive Statistics

| variable | description | mean | sd | min | max |
|---|---|---|---|---|---|
| age | in years | 44.31 | 10.75 | 21 | 76 |
| education | In stages from 1 (=none) to 6 (=university) | 2.70 | 1.19 | 0 | 7 |
| exporting | Whether the farmer produces pineapples for export | .27 | .44 | 0 | 1 |
| price | Pineapple price, in New Ghanaian Cedis (GHC) | .41 | .16 | 0 | .7 |
| income | Annually from pineapple, in New Ghanaian Cedis (GHC) | 2171 | 3599 | 0 | 27k |
| land | Farmland that is used for pineapples, incl. fallow and rotations | 3.90 | 5.22 | .5 | 50 |
| tenure | Perceived tenure security on a scale of 1 (highly insecure) to 6 (secure) | 5.52 | 1.09 | 0 | 6 |
| land rented | Percentage of land rented for growing pineapples | .70 | .45 | 0 | 1 |
| loan | Whether the farmer has access to credit | .33 | .47 | 0 | 1 |
| nonfarm inc | Whether the farmer receives non-farm income | .21 | .41 | 0 | 1 |
| new variety | Whether the new MD2 variety is grown | .31 | .46 | 0 | 1 |
| residuals | Whether crop residuals are used | .07 | .26 | 0 | 1 |
| mulching | Whether mulching is used | .53 | .49 | 0 | 1 |
| fertilizer | Whether chemical fertilizer is used | .52 | .49 | 0 | 1 |
| farming skills | Factor variable from years of farming experience, years of Pineapple experience, and self-categorization | 5.32 | 1.29 | -4.10 | 0.51 |
| start f | First year the farmer started own farm | 1999 | 8.91 | 1962 | 2013 |
| start p | First year the farmer started growing pineapples | 2002 | 7.14 | 1978 | 2013 |
| expertise | Perception, from 1 (=less experience than peers) to 3 (=more) | 1.71 | .73 | 0 | 4 |
| social capital | Factor variable from frequency of social events, number of people who would lend farmer money, and generalized trust | 0 | .35 | -1.97 | 6.32 |
| social events | Reported attendance of social events from 1 (=never) to 6 (=often) | 4.43 | 1.78 | 0 | 6 |
| borrow | Number of people the farmer could borrow money from | 1.83 | 2.91 | 0 | 30 |
| trust | How much the farmer generally trusts others, from 1 (=you cannot be too careful) to 6 (=most people can be trusted) | 2.55 | 1.83 | 0 | 6 |
| leader ideas | How open are local opinion leaders for new ideas, from 1 (=not at all) to 6 (=very much) | 5.25 | 1.17 | 0 | 6 |
| leader trad | How traditional are local opinion leaders, as above | 3.80 | 1.66 | 0 | 6 |
| leader innov. | How innovative are local opinion leaders, as above | 5.19 | 1.21 | 0 | 6 |
| Accra dist. | Distance between farms and the capital | 53Km | 38Km | 5 | 135Km |
| company dist. | Distance between farms and processing companies | 46Km | 35Km | 3 | 125Km |
| coast dist. | Distance between regions and coast | 23Km | 32Km | 873 | 133Km |
| cereals | Historical reliance of a region on cereals (%) | .07 | .25 | 0 | 1 |
| roots | Historical reliance of a region on roots and tubers (%) | .62 | .49 | 0 | 1 |
| suit. cereals | Soil suitability for this kind of crop, from 1 (=not) to 7 (=optimal) | 3.64 | 1.46 | 1 | 6 |
| suit. roots | Soil suitability for this kind of crop, from 1 (=not) to 7 (=optimal) | 4.55 | 2.00 | 1 | 7 |
| slavery impact | Number of slaves taken per region, $15^{th}$ - $18^{th}$ century | 16675 | 5386 | 10354 | 21485 |
| rain | Reported local rainfall, from 1 (=very bad) to 6 (=optimal) | 4.48 | 1.32 | 1 | 6 |
| soil | Reported, from 1 (=not a problem) to 4 (=big problem) | 1.65 | .78 | 1 | 4 |
| elevation | Calculated in ArcGIS, in meters | 85.15 | 61.62 | 9 | 2.574 |
| topography | Calculated in ArcGIS, standard deviation in meters | 42.16 | 38.26 | 587 | 1.553 |
| m. schools | Number of colonial missionary schools in the region of the farms | 16.24 | 11.68 | 1 | 48 |

On average 7% of the farmers have ancestors who relied predominantly on cereal farming (mainly maize), which was approximately located in an area north of Accra. It can also be seen in table 1, that on average several thousand slaves were taken from the ancestral communities of Today's pineapple farmers.

Our empirical framework is simple. First, we estimate the effect of historical experiences on current farm incomes:

$$Y_{ijk} = \alpha_{jk} + \beta_1 H_{ijk} + \beta_2 X_{ijk} + \varepsilon_{ijk} \qquad (1)$$

where $Y_{ijk}$ is the income of farmer $i$ of ethnicity $i$ in farm group $k$, $\alpha_{jk}$ are fixed effects, $X_{ijk}$ are control variables and $H_{ijk}$ are historic experiences. Our control variables include age, education and farming skills, farm size, soil suitability, rainfall, topography, distance to Accra, and the pineapple price. Tenure security is not included because the farmers use relatively safe land for their pineapple production, so there is little explanatory power in this variable.

The age of the farmers is measured in years, their education level on a scale from 1 (=none) to 6 (Tertiary/University) plus a seventh option (=other). Farming skills are obviously difficult to measure but we use a factor variable from the years of farming experience in general and with pineapples and a self-report. For the farm size we use the land that can be used to grow pineapples in hectares. This includes land that is not currently used (e.g. fallow) but it does not include shared land under traditional land rights, as this land is commonly perceived as not sufficiently secure to grow a cash crop. Because farmsize is self-reported it is often rather guessed than known by the farmers and should thus be seen as a proxy. Soil suitability captures the soil suitability for pineapple and is measured on a scale from 1 (=no constraint) to 4 (= big constraint). Rainfall is the reported rainfall quantity from (1= very bad) to 6 (= optimal). Generally, reported rainfall agrees with rainfall measured at weather stations, but weather station data is very coarse (there are only 3 weather stations four our whole sample) and the effect of rainfall is highly dependent on micro-climate and soil. To capture effects of the topography, we use Arcgis software to calculate the mean elevation of the farms and its standard deviation. Finally, we control for the distance to Accra in Kilometers, also computed in Arcgis software.

Testing the correlation amongst our explanatory variables, we find that most variables only correlate weakly with each other, except for a rather high correlation between the historic slavery impact and whether or not the farmer now grows the modern MD2 variety (-.58) and that the descendants of cereal farmers live in more rugged areas (+.42).

We use (1) to estimate the income of the farmers and use the difference between actual and predicted income as left hand side in our main analysis:

$$Y_{ijk} - \hat{Y}_{ijk} = \alpha_{jk} + \beta_1 S_{ijk} + \beta_2 X_{ijk} + \varepsilon_{ijk} \quad (2)$$

where $S_{ijk}$ is the explanatory variables that is hypothesized to explain the income difference.

To take into account unobserved heterogeneity, we estimate 2SLS regressions:

$$\begin{gathered} Y_{ijk} - \hat{Y}_{ijk} = \alpha_{jk} + \beta_1 \hat{S}_{ijk} + \beta_2 X_{ijk} + \varepsilon_{ijk} \\ S_{ijk} = \alpha_{jk} + \beta_1 Z_{ijk} + \beta_2 X_{ijk} + \varepsilon_{ijk} \end{gathered} \quad (3)$$

To be specific, in equation (1) we use two historic variables that have been demonstrated to affect current incomes. The first is the long-term impact of the trans-Atlantic slave trade (Nunn 2008) and the other is the long-term impact of having ancestors who depended on different kinds of crops (Wuepper and Drosten 2016). Both variables have changed the trajectory of cultural evolution – the former affected social capital and the second affected the belief of the farmers to be able to profit from agricultural investments ("investment self-efficacy"). Of course, not everybody is equally affected by history and there are farmers whose ancestors where highly affected by the trans-Atlantic slave trade and whose ancestors also depended on the "wrong kind of crop" and who nevertheless achieve a good income. We test the explanatory power of human and social capital. We are carefully considering the possible endogeneity of our explanatory variables and use 2SLS regressions. For the participation in the pineapple export, we use as instrument the distance between the farms and the exporting companies. This variable predicts the participation but has no other income effect, because the companies are sufficiently far away from the coast and the major cities Accra and Cape Coast. Human capital as proxied by the education level of the farmers can be taken as exogenous variable because education is completed before income starts. Social capital, on the other hand, could be endogenous. To proxy social capital, we use a factor variable from how much the farmers trust others, how many would borrow them money if needed, and how frequently they attend social events. We also use the last variable as instrument for social capital, which means we only use the commonality of trust, borrowers, and social interaction that is purely explained by the frequency of social interaction. Social interaction has no direct effect on income, but it is established in the literature that social interaction increases social capital (Feigenberg et al. 2013).

The approach described above is not without its critics. The criticism concerns equation (1) in which a variable from one period is used to explain an outcome from another period. Austin (2008) calls this a “Compression of History” because all the other periods are ignored. Using longitudinal data has the clear advantages that it can be seen whether the effect of a historical variable is stable in time, or whether it changes and the danger of picking up spurious effects is also reduced. However, it would be a mistake to right away discard all empirical evidence using data from only two periods. As Fenske (2011) argues, the new economic history is characterized by its careful focus on causal inference, which we argue does not require longitudinal data but a sound research design using the tools of modern econometrics (Angrist and Pischke 2008). In the following, we consider two historic experiences that changed the trajectories of cultural evolution in rural Ghana. Culture takes a long time to develop and is both a cause and a product of the context (Boyd and Richerson 1985, Boyd et al. 2011). Thus, even though the long time span in our data means that actors and context are far from stable, younger generations are still connected with the older generations, because cultural evolution means that they inherit heuristics and other information as a function of history.

## 3. Predicting Farm Incomes with History

First of all, we must establish that generally, both historic variables have a significant effect on current farm incomes.

**Table 3. Statistical Relationships between Historical Variables and Current Farm incomes**

| | (1) | (2) |
|---|---|---|
| dep var | income | income |
| model | OLS | OLS |
| cereals | 0.0646** | |
| | (0.0256) | |
| slaves | | -0.306*** |
| | | (0.0221) |
| land | 0.106*** | 0.108*** |
| | (0.0389) | (0.0397) |
| loan | 0.0691** | 0.0667** |
| | (0.0307) | (0.0324) |
| education | 0.0439*** | 0.0464*** |
| | (0.0154) | (0.0153) |
| controls | yes | yes |
| FE(district and ethnicity) | yes | yes |
| R-sq | 0.289 | 0.296 |
| N | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at the district and the ethnicity level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***). We control for the pineapple suitability of the soil, rainfall, topography (elevation and ruggedness), distance to Accra, farmsize, credit access, pineapple price, age, education, and whether the farmer's descendant were cereal farmers and how much they were impacted by the external slave trades.

Because most of the farmers do not reliably take accounts of their revenues and costs, there is significant measurement error in their reported incomes. Our approach to make the income variable more precise is to estimate a factor variable, with calculated incomes from detailed statements on all in- and outputs, costs and prices on the one hand, and stated income categories on the other. We interpret the commonality of both income-measures as better approximation to actual incomes than either variable alone.

To establish that a historical dependency on cereals and the fraction of slaves taken have significant long term effects in our sample, we first show the basic relationships with OLS regressions below (table 3) and subsequently estimate 2SLS regressions to establish that the observed relationships are causal (table 4).

Table 3 shows that farmers with cereal farming ancestors have higher incomes from pineapples than others and the same is true for farmers whose ancestors were less impacted by the trans-Atlantic slave trade. We also show the estimated relationships with three selected current variables, to give an impression of scales. Full results of all estimations can be found in the appendix.

To investigate whether these correlations indicate causal effects of historic events on current farm incomes, our identification strategy follows that of Wuepper and Drosten (2016) and Nunn and Wantchekon (2011).

The reason why historic production systems still affect economic outcomes is because the experienced returns on investments (by the farmers' ancestors) shaped different "investment-cultures". In regions where cereals where grown, farmers had higher incentives to invest and learned about the profitability of such investments. In regions where roots, tubers, or trees were the main subsistence-basis, the farmers learned that investments are far less profitable. These cultural beliefs are highly persistent, because they are self-reinforcing. Farmers who belief investment are not profitable are less likely to profit, because they often invest too little and withdraw to quickly. The difficulty in identifying the long-term effect of growing cereals versus other crops is that it is possible that ancestral communities self-sorted into different production systems because they already had different beliefs about the profitability of agricultural investments. To avoid this possible endogeneity bias, we use the regional soil suitability for different kinds of crops as instruments (the soil suitability data can be obtained from FAO's GAEZ database). Specifically, we use the regional suitability for maize as a cereal and yams as a root. As can be seen in table 3, regions that were historically more suitable for growing cereals (maize) are indeed more likely to have historically depended on cereal-farming (than other regions, that where more suitable to grow other crops, such as yams or other roots and tubers). The exclusion restriction plausibly holds because the farmers cannot have changed their regional soil suitability in order to grow their preferred crop. It is of course possible that historically farmers migrated to the areas where their preferred crop could be grown. This kind of selection effect however, does not cause endogeneity, because this would still imply that Ghana's regional variation in "investment-self-efficacy" stems from the different historic production systems. Table 3 indicates that this has a significant impact on current farm incomes.

The reason why the experience of the trans-Atlantic slave trade has such a strong, long-term effect on economic development in Africa is because it eroded inter-personal trust, which is commonly seen as an important part of social capital. As research on very different levels has shown, social capital is quite important for the economy, as investigated by Knack and Keefer (1997), Fukuyama (2001), Ahlerup et al. (2009), or Karlan (2005) and Feigenberg et al. (2010). For Ghana's farmers, social capital is often a substitute for formal insurance schemes and a source of credit and information.

To connect Ghana's 2013 pineapple farmers with historic data on slave exports, we use the data of Nunn (2008), Nunn and Wantchekon (2011) and Murdock (1967) to overlay the spatial distribution of the slavery impact and the spatial distribution of our sampled farmers. Thus, farmers and slavery impact are joined on a geographical basis. (We also tested joining the variables based on reported ethnicities but we do not obtain enough variation within the south of Ghana to get statistically significant results.)

The difficulty of identifying the causal effect of the trans-Atlantic slave trade on current farm incomes is – again – that societies could have self-selected because they already had distinct levels of trust. To avoid this endogeneity-bias, we use the distance between the regions and the coast as an instrument. As Nunn and Wantchekon (2011) establish, the coast distance had a significant effect on the slave trades because of the implied transaction costs (mainly the transport of the slaves) but not directly on trust, as indicated by an absence of a relationship between coast-distance and trust on continents without the major slave trades.

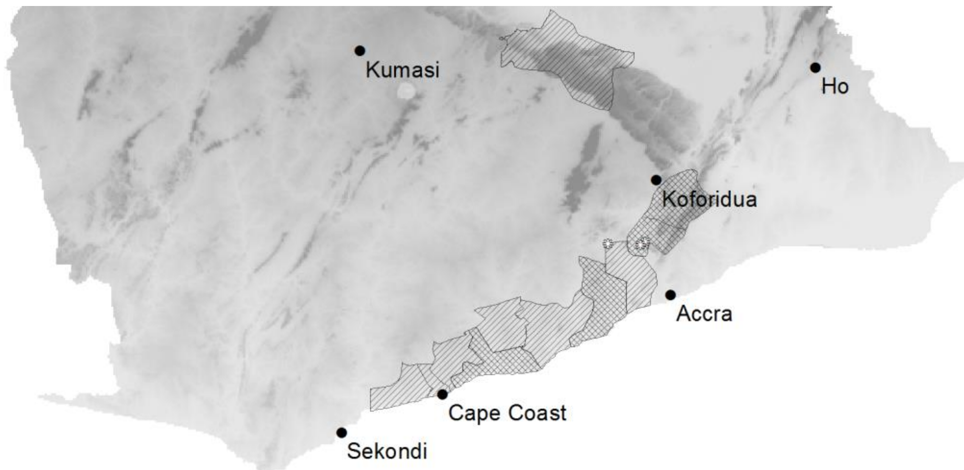
**Table 4. Historically Predicted Incomes (2SLS)**

| RF | (1) | (2) |
|---|---|---|
| dep.var. | income | income |
| model | 2SLS | 2SLS |
| cereals | 0.292** | |
| | (0.130) | |
| slavery | | -0.208*** |
| | | (0.130) |
| 1ST | cereals | slavery |
| suitability cereals | 0.0701* | |
| | (0.0392) | |
| coast distance | | -0.442*** |
| | | (0.0156) |
| controls | yes | yes |
| FE (districts) | yes | yes |
| F excluded | 11.20 | 803.68 |
| R-sq 2nd stage | 0.10 | 0.16 |
| R-sq 1st stage | 0.92 | 0.99 |
| N | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at the district and the ethnicity level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).We control for the pineapple suitability of the soil, rainfall, topography (elevation and ruggedness), distance to Accra, farmsize, credit access, pineapple price, age, education, and whether the farmer's descendant were cereal farmers and how much they were impacted by the external slave trades.

In our case, the exclusion restriction only holds if coast distance does not have a direct effect on farm incomes. Because we control for the distance to the capital, as well as rainfall, soils, elevation and topography, this assumption should hold, as these variables are the plausible channels for any direct effect of coast distance. It can be seen in table 3, that coast distance reduced the impact of the trans-Atlantic slave trade and this in turn increases current farm incomes.

**Figure 1. Incomes Predicted by History**

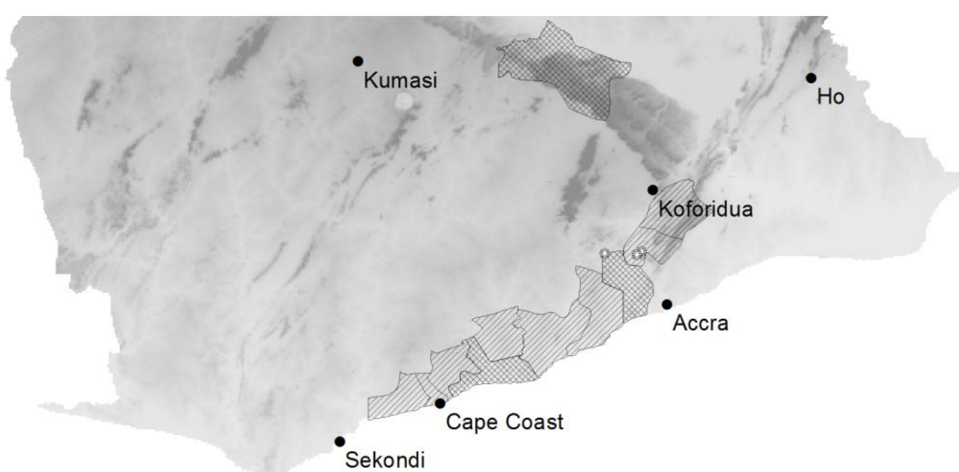



Notes: Underlying map shows the topography of Ghana's south (darker areas are more elevated). Striped regions have lower and squared regions have higher predicted farm incomes, based on historic experiences with slavery and agricultural production as well as biogeographic variables.

We use specifications (1) and (2) to predict the incomes of the farmers. Figure 1 shows a map of southern Ghana and the spatial distribution of historically predicted incomes, conditional on controlling for bio-geographic variables and based on both the impact of the trans-Atlantic slave trade and the historical farming systems.

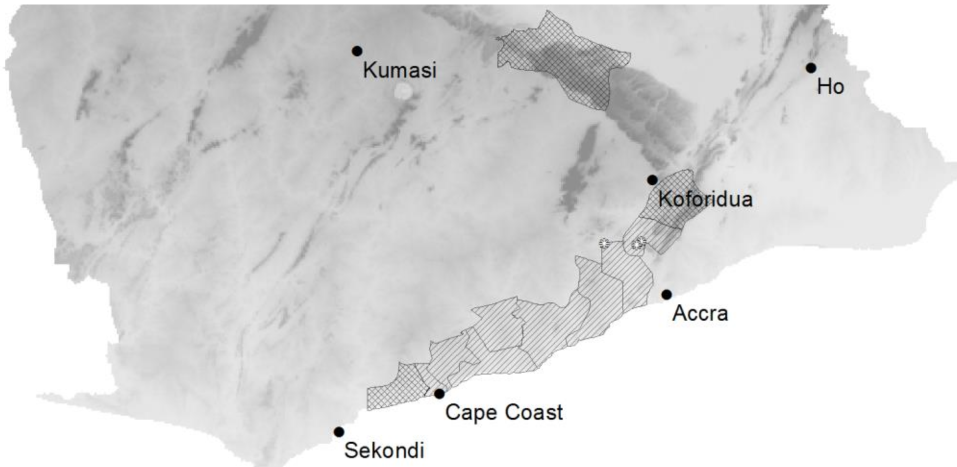
Figure 2 and 3 show where in Ghana farmer achieve higher than predicted incomes. This of course masks immense heterogeneity within the regions but hopefully helps to communicate the basic idea of our research.

**Figure 2. Improvements Relative to Predicted Incomes from Historical Production System**




Notes: Squared regions achieve on average higher incomes than predicted by their ancestral farming systems, whereas striped regions do not.

**Figure 3. Improvements Relative to Predicted Incomes from Trans-Atlantic Slave Trade**




Notes: Squared regions achieve on average higher incomes than predicted from their ancestral slavery impact, whereas striped regions do not.

As an example, the region furthest north on the map is Kwahu South. It is located in relatively high altitude and as shown in figure 1, based on history and biogeography, has predicted farm incomes that are rather low. However, in figures 2 and 3, it can be seen that many farmers achieve better incomes than predicted, so perhaps we can learn how they did it.

This example is also useful to see a second point: Despite the observation that many farmers in Kwahu achieve higher incomes than predicted, their absolute income still remains lower than that in most other regions in southern Ghana. All we find is that it is higher compared to what it would be if history would fully determine incomes. Overall, there is a positive relationship between beating the predictions and farm incomes.

## 4. Who Achieves a Higher Income?

We begin by testing for a statistical relationship between the way the farmers produce and market their pineapples on the one hand, and whether they top their historically predicted income on the other (table 5). We estimate two specifications. The first dependent variable is the difference between the actual income a farmer achieves and her predicted income from how much her ancestors where impacted by the trans-Atlantic slave trade (Difference1). The second dependent variable it the difference between the actual income a farmer achieves and her predicted income from which historical farming system her ancestors depended on (Difference2).

**Table 5. OLS Regressions on Income Improvements (A)**

| | (1) | (2) |
|---|---|---|
| dep.var. | Difference1 | Difference2 |
| model | OLS | OLS |
| export | 0.0630** | 0.0643** |
| | (0.0258) | (0.0284) |
| new variety | 0.0946*** | 0.0856** |
| | (0.0240) | (0.0368) |
| residues | 0.0701*** | 0.0681*** |
| | (0.0110) | (0.0123) |
| mulching | 0.0322*** | 0.0327*** |
| | (0.0105) | (0.0114) |
| fertilizer | 0.0898** | 0.0825** |
| | (0.0395) | (0.0407) |
| controls | yes | yes |
| FE (districts, ethnicity, groups) | yes | yes |
| R-sq 1st stage | 0.36 | 0.43 |
| N | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects.

Independent from the historical variable that was used to predict incomes, those who do better are likely to produce for export and in a more intensive way (growing the modern MD2 variety and using chemical fertilizer, mulching, crop residues, and fertilizer). Throughout, we control for district, ethnicity, and farm group fixed effects and standard errors are clustered at the farm group and ethnicity level. We also always include a vector of individual controls, including bio-geographic variables and socio-economic characteristics of the farmers.

An obvious shortcoming of table 5 is that it shows *what* farmers do who achieve higher than predicted incomes, but not *why*. Also, we must address the issue of endogeneity again, as the indicated relationships might not be causal.

We start with table 6, where we investigate whether opinion leaders, education level and social capital are significantly correlated with higher than predicted incomes. Indeed, farmers who report that the opinion leaders in their communities are more open for new ideas and less traditional, are more likely to beat their historical predictions. Similarly, education and social capital both increase the probability that a farmer has a better than predicted income.

**Table 6. OLS Regressions on income Improvements (B)**

| | (1) | (2) |
|---|---|---|
| dep.var. | Difference1 | Difference2 |
| model | OLS | OLS |
| leader ideas | 0.0462** | 0.0461** |
| | (0.0192) | (0.0209) |
| leader traditional | -0.0415** | -0.0386** |
| | (0.0188) | (0.0191) |
| social capital | 0.0634* | 0.0643* |
| | (0.0339) | (0.0336) |
| education | 0.0595*** | 0.0576*** |
| | (0.0107) | (0.00992) |
| controls | yes | yes |
| FE (districts, ethnicity, groups) | yes | yes |
| R-sq | 0.32 | 0.40 |
| N | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects.

There are different reasons why we might want to base our analysis on an instrumental variables framework (table 7). First of all to understand whether exporting really helps farmers to achieve higher than predicted incomes depends on the previous selection of who gets to export (Bigsten et al. 2000, Van Biesebroeck 2005, Wagner 2007, Harou and Walker 2012). We use distance to the next exporting company as instrument, arguing that because of transaction costs, distance greatly matters, and because the companies are located sufficiently far away from a major city, the coast and other confounding locations, the company distance does not violate the exclusion restriction. Notably, adjusting for the fact that companies select better farmer and better farmers self-select, there is no causal effect of producing for export and achieving a higher than predicted income.

One might also worry that social capital is in part an outcome of economic success, if trust and borrowing money are used as indicators. Thus, we use the attendance of social events as instrument, as Feigenberg et al. (2010) show that simply increasing social interactions can build social capital, and attending weddings and funerals is unlikely to have any direct effect on farm incomes. The result is a strengthening of the estimated effect of social capital, because the instrumented variable is a more precise measure than the initial factor variable, which is more influenced by measurement error.

Estimating the effect of the local opinion leaders poses a similar identification challenge as the effects of exporting and social capital did. The challenge of identifying the opinion leader effect is the "reflection problem" discussed by Manski (1993, 2000). As with a person standing in front of a mirror, it is difficult to figure out, whether the mirror reflects the person or the person reflects the mirror, as both move simultaneously. In our example, it could be the case that innovative opinion leaders cause others to follow their steps. However, it could also be the case that all farmers, including the opinion leaders, react to the same incentives and constraints and thus behave similarly.

**Table 7. 2SLS Regressions on Income-Improvements**

| RF | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| dep.var | Difference1 | | Difference2 | |
| model | 2SLS | 2SLS | 2SLS | 2SLS |
| export | 0.0443<br>(0.0604) | | 0.104<br>(0.0726) | |
| social capital | | 0.115***<br>(0.0201) | | 0.115***<br>(0.0187) |
| education | 0.0571***<br>(0.0131) | | 0.0472***<br>(0.0109) | |
| 1ST | export | social capital | export | social capital |
| comp. dist. | -.594***<br>(0.183) | | -.594***<br>(0.183) | |
| social events | | 0.761***<br>(0.038) | | 0.761***<br>(0.038) |
| controls | yes | yes | yes | yes |
| FE districts, ethnicity and groups | yes | yes | yes | yes |
| R-sq (2nd stage) | 0.30 | 0.30 | 0.31 | 0.31 |
| R-sq (1st stage) | 0.6 | 0.7 | 0.6 | 0.7 |
| F excluded 1ST | 10.68 | 201.62 | 10.68 | 201.62 |
| N | 398 | 398 | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects. F excluded Is the F-value for the excluded instrument(s).

It is generally difficult to find a feasible instrument that explains why the opinion leaders in a community are innovative and that does not affect anybody else. It is easier to find an instrument that enables us to perform a falsification test (table 8).

**Table 8. Falsification Test for Leadership Effect**

| RF | (1) | (2) |
|---|---|---|
| dep.var. | Difference1 | Difference2 |
| model | 2SLS | 2SLS |
| leader idea | 0.0888<br>(0.0713) | 0.0827<br>(0.0718) |
| 1ST | leader idea | leader idea |
| neighbors idea | 0.478***<br>(0.091) | 0.548***<br>(0.098) |
| controls | yes | yes |
| FE (districts, groups and ethnicity) | yes | yes |
| R-sq (2nd stage) | 0.29 | 0.37 |
| R-sq (1st stage) | 0.16 | 0.16 |
| F excluded | 27.45 | 31.22 |
| N | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects. F excluded Is the F-value for the excluded instrument(s).

We use the fact that incentives and constraints a spatially continuously distributed (rainfall, soil quality, prices, credits are all commonly a function of the distance to some sort of center, such as a gradient from west to east, distance to a city, or distance to a bank). The opinion leaders, in contrast, are located within communities and farm groups, and thus not continuously distributed.

**Table 9. Further Investigations**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| dep.var | Difference1 | | Difference2 | |
| model | 2SLS | 2SLS | 2SLS | 2SLS |
| social capital (SC) | 0.0982*** | 0.109*** | 0.0971*** | 0.108*** |
| | (0.0264) | (0.0284) | (0.0267) | (0.0288) |
| leader ideas (L) | 0.0129 | 0.0252 | 0.00794 | 0.0204 |
| | (0.0200) | (0.0244) | (0.0200) | (0.0246) |
| nonfarm_inc | 0.175*** | | 0.176*** | |
| | (0.0501) | | (0.0498) | |
| land | 0.0835*** | | 0.0854*** | |
| | (0.0240) | | (0.0245) | |
| export | 0.0293 | | 0.0306 | |
| | (0.0340) | | (0.0342) | |
| loan | 0.0887*** | | 0.0935*** | |
| | (0.0256) | | (0.0256) | |
| education | 0.0332** | 0.0586*** | 0.0304** | 0.0566*** |
| | (0.0160) | (0.0161) | (0.0154) | (0.0156) |
| age | -0.00849 | -0.00441 | (0.0161) | -0.0134 |
| | (0.0164) | (0.0185) | -0.0163 | (0.0134) |
| price | -0.0168 | -0.0141 | -0.0163 | -0.0134 |
| | (0.0119) | (0.0132) | (0.0120) | (0.0134) |
| farming skills | -0.0118 | -0.00843 | -0.00773 | (0.0194) |
| | (0.0192) | (0.0189) | (0.0195) | -0.0508*** |
| Biogeography | yes | yes | yes | yes |
| FE districts and ethnicity | yes | yes | yes | yes |
| FE farm group | yes | yes | yes | yes |
| F excluded SC | 96.88 | 192.26 | 84.15 | 192.26 |
| F excluded L | 20.03 | 20.36 | 39.24 | 20.36 |
| R-sq (1ST) SC | 0.67 | 0.66 | 0.71 | 0.66 |
| R-sq (1ST) L | 0.61 | 0.61 | 0.66 | 0.61 |
| R-sq (2nd) | 0.43 | 0.30 | 0.50 | 0.39 |
| N | 398 | 398 | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects. F excluded Is the F-value for the excluded instrument(s).

By instrumenting the innovativeness of the opinion leaders with the innovativeness of adjacent communities, we predict how innovative the opinion leaders would be, purely based on regional incentives and constraints. As can be seen in table 8, the innovativeness of the opinion leaders in adjacent communities has indeed strong predictive power for their neighboring opinion leaders, but the instrumented variable does not have a significant income effect.

Table 9 presents the second stages of further 2SLS regressions. We omit the first stages for brevity but report $R^2$ and F test of the excluded instruments. We use the gained space to also present the estimates for our control variables, except fixed effects and biogeography.

It can be seen that the estimated effects of social and human capital are robust to the inclusion of different control variables.

**Table 10: Historic Roots of Social Capital**

| RF | (1) | (2) |
|---|---|---|
| Table Dep.var | Difference1 | Difference2 |
| model | 2SLS | 2SLS |
| social capital | 0.112*** | 0.111*** |
| | (0.0306) | (0.0331) |
| 1ST | social capital | social capital |
| missionary schools | -0.210*** | -0.210*** |
| | (0.0481) | (0.0481) |
| social events | 0.761*** | 0.761*** |
| | (0.0417) | (0.0417) |
| controls | yes | yes |
| FE (districts) | yes | yes |
| R-sq RF | 0.30 | 0.38 |
| R-sq 1ST | 0.66 | 0.66 |
| Sargan P | 0.74 | 0.74 |
| F excluded | 152.69 | 152.69 |
| N | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects. F excluded Is the F-value for the excluded instrument(s).

A remaining concern is that social capital might also be the outcome of a historical experience, and thus not actually a factor that helps to "overcome" history. This concern is raised by the research of Wuepper and Sauer (2016), who find that especially during Ghana' colonial times, Christian missionary schools often tried to separate their new converts from their old communities (to avoid their re-conversion) and that such a strategy persistently lowered social capital in the targeted communities.

To compare how much social capital is explained by current social events and how much is explained by Christian missionary activity, we use both variables as instruments and estimate once more, who achieves a higher than predicted income (table 10). Of interest are the first stages. It can be seen that indeed, the Christian missionary schools predict a lower level of social capital today, but the estimated effect of this variable is less than a third of the estimated effect of the current event variable.

## 5. Robustness Analysis

As a first robustness check, we re-estimate the specifications from table 7 but exclude Kwahu South (table 11). As can be seen on the maps in section 3, Kwahu is somewhat different from the other sampling regions. Especially, it is in a more mountainous region and relatively remote. Nunn and Puga (2012) find that communities in more rugged terrain were more protected against slavery, thus we might have overestimated the slavery impact based on relatively coarse statistics, which could produce an upwards bias in the income improvements of the farmers in Kwahu. Furthermore, e.g. Barrett et al. (2012) emphasize the importance of transaction costs for agricultural contracts, so that Kwahu's distance to markets and companies could possibly introduce a discontinuity in the chance to produce for export. For these reasons, table 10 must establish that the results are robust to the exclusion of potential outlier farmers. Indeed, the estimated effects do not change substantially between table 7 and 11.

**Table 11. Excluding Kwahu South**

| RF | (1) | (2) |
|---|---|---|
| dep.var. | Difference1 | Difference2 |
| model | 2SLS | 2SLS |
| social capital | 0.120*** | 0.118*** |
| | (0.0311) | (0.0311) |
| education | 0.0635*** | 0.0618*** |
| | (0.0162) | (0.0157) |
| 1ST | social capital | social capital |
| social events | 0.770*** | 0.770*** |
| | (0.0534) | (0.0534) |
| controls | yes | yes |
| FE (districts) | yes | yes |
| R-sq (2nd stage) | 0.30 | 0.38 |
| R-sq (1st stage) | 0.657 | 0.657 |
| F excluded | 196.68 | 196.68 |
| N | 371 | 371 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects. F excluded Is the F-value for the excluded instrument(s).

A second robustness check is to exclude the descendants of cereal farmers and to test whether the historic suitability to grow cereals explains the income of the remaining farmers. This test is motivated by the fact that only a small share of our sampled farmers actually descents from cereal farmers, so we need to test see whether cereal suitability explains current farm incomes through other channels than the historical choice of production systems. Table 12 suggests that cereal suitability only matters for the descendants of cereal farmers, which increases our confidence in the instrument.

**Table 12. Excluding Farmers whose Ancestors were Cereal Farmers**

| | |
|---|---|
| dep var | income |
| model | OLS |
| sample | Without descendants of cereal farmers |
| cereals | -0.0621 |
| | (0.0400) |
| land | 0.0952** |
| | (0.0375) |
| loan | 0.0731* |
| | (0.0381) |
| education | 0.0473** |
| | (0.0216) |
| controls | yes |
| FE(district and ethnicity) | yes |
| R-sq | 0.34 |
| N | 272 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital. FE is short for fixed effects.

**Table 13. Testing for Omitted Variables**

| spec. | (1) | (2) |
|---|---|---|
| dep.var | Difference1 | Difference2 |
| model | OLS | OLS |
| slavery | -0.158 | |
| | (0.104) | |
| cereals | | 0.0711 |
| | | (0.101) |
| controls | yes | yes |
| FE(district, group, and ethnicity) | yes | yes |
| R-sq | 0.268 | 0.269 |
| N | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education, farming skills, farmsize, pineapple price, credit access, rain, soil suitability, topography, and distance to the capital.

A third robustness check is to include the impact of the trans-Atlantic slave trade and the historical dependency on cereals as explanatory variables for the income differences. As tables 13 shows, they are not significant.

Because Huillery (2009, 2011) and Jedwab and Moradi (2015) find that colonial investments often continue to explain economic outcomes, we must also test whether they do so in our sample. In table 14, we regress the income of the farmers on their distance to the colonial railroad tracks and the amount of colonial government schools in their community. Neither the colonial investments into infrastructure nor education explains income differences amongst the pineapple farmers.

**Table 14. The Effect of Colonial Investments on Current Farm Income**

| spec. | (1) |
|---|---|
| dep.var. | income |
| model | OLS |
| government schools | 0.0105 |
| | (0.0214) |
| railroad distance | 0.0339 |
| | (0.0891) |
| controls | yes |
| FE(district and ethnicity) | yes |
| R-sq | 0.16 |
| N | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at the district and the ethnicity level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***). In addition to our usual set of control variables, we also control for the coast distance because the government schools where not randomly placed but only close to the coast, near Cape Coast and Accra.

Finally, an important institutional legacy could be land tenure. We have argued already above that pineapple is predominantly grown on relatively safe, rented land, and not on land under traditional, less secure land. As a final test, we show in table 15 the results of two specifications. In the first, we regress the income of the farmers on their reported tenure security and in the second, we replace this variable with the share of rented land that is used to produce the pineapples. Neither variable is significant.

**Table 15. The Role of Tenure Security**

| spec. | (1) | (2) |
|---|---|---|
| dep.var. | income | income |
| model | OLS | OLS |
| tenure security | 0.0149 (0.0114) | |
| land rented | | 0.0251 (0.0478) |
| Controls | | |
| FE (districts and ethnicity) | | |
| R-sq | 0.154 | 0.153 |
| N | 398 | 398 |

Notes: Table shows estimated coefficients and standard errors (in brackets). Standard errors are clustered at the district and the ethnicity level. Significance levels are 0.1 (*), 0.05 (**), and 0.01 (***).

In summary, we find that famers who achieve better than predicted incomes farm more intensively, export their produce and have more innovative opinion leaders, more education, and higher social capital. A robust causal effect is only found for social and human capital. Export production does not have any causal effect, as any significant income differences are caused by a selection effect. The effect of more open and innovative opinion leaders also does not seem to be causal, as we cannot rule out that some farm groups have generally more innovative and open farmers, including their opinion leaders and there is no significant causal effect of the opinion leaders on their peers. We find that the opinion leaders lose their significance once we control for farm group fixed effects, credit access, non-farm income and farmland. We cannot tell whether this is because these are outcomes of having certain opinion leaders around or whether these variable are merely correlated with the innovativeness of the farmers and thus create a spurious effect.

We also find that farmers who achieve higher than predicted incomes farm more intensively. This could be a channels through which social capital and education materialize as financial gain.

To formally investigate this, we present the results of a few simple mediation analyses in table 16 for social capital and in table 17 for education.

**Table 16. Mediation Analysis for Social Capital**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | Difference1 | | | Difference2 | |
| model | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS |
| social capital | 0.0300* | 0.0310* | 0.0404** | 0.0387** | 0.0402** | 0.0507*** |
| | (0.0156) | (0.0186) | (0.0170) | (0.0163) | (0.0201) | (0.0181) |
| production | 0.294*** | | | 0.305*** | | |
| | (0.0663) | | | (0.0774) | | |
| fertilizer | | 0.326*** | | | 0.330*** | |
| | | (0.0629) | | | (0.0782) | |
| export | | | 0.0919*** | | | 0.0847** |
| | | | (0.0289) | | | (0.0325) |
| ACME | .02*** | .02*** | .01*** | .02*** | .02*** | .01*** |
| Direct Effect | .03*** | .03*** | .04*** | .04*** | .04*** | .05*** |
| Total Effect | .05*** | .05*** | .05*** | .06*** | .06*** | .06*** |
| % of Tot Eff mediated | .41*** | .38*** | .20*** | .36*** | .33*** | .20*** |
| R-sq | 0.17 | 0.15 | 0.09 | 0.17 | 0.14 | 0.15 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, education , rain, soil suitability, topography, and distance to the capital. ACME stands for average causal mediation effects

As before, the dependent variable in the first three specifications is the difference between actual farm income and the income predicted by the slavery impact and the dependent variable in the last three specifications is the difference between farm income and the income predicted by the historical farming systems. We create two factor variables to test for causal channels of social capital on income. The first factor variable is production, which has exporting, mulching, the grown variety, fertilizer and the use of residues as inputs. The second factor variable only has fertilizer, the grown variety and the use of crop residues as inputs. Table 16 suggests that about 30% of the effect of social capital on farm incomes is mediated through the decisions which variety to grow and whether to use chemical fertilizer and crop residues. Between 15 and 15% of the effect might be mediated through the decision to produce for export.

As can be seen in table 17, less than 10% of the education effect is mediated through the decisions which variety to grow and whether to use chemical fertilizer and crop residues.

**Table 17. Mediation Analysis for Education**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | Difference1 | | | Difference2 | |
| model | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS | OLS/OLS |
| education | 0.0505*** | 0.0565*** | 0.0507*** | 0.0536*** | 0.0597*** | 0.0543*** |
| | (0.0180) | (0.0155) | (0.0182) | (0.0190) | (0.0162) | (0.0199) |
| production | 0.294*** | | | 0.305*** | | |
| | (0.0663) | | | (0.0774) | | |
| fertilizer | | 0.326*** | | | 0.330*** | |
| | | (0.0629) | | | (0.0782) | |
| export | | | 0.0919*** | | | 0.0847** |
| | | | (0.0289) | | | (0.0325) |
| ACME | .004* | .001 | .004* | .004* | .001* | .003* |
| Direct Effect | .050*** | .060*** | .050*** | .053*** | .059*** | .054*** |
| Total Effect | .054*** | .055*** | .054*** | .057*** | .058*** | .057*** |
| % of Tot Eff mediated | .08*** | .03*** | .07*** | .07*** | .02*** | .06*** |
| R-sq | 0.17 | 0.15 | 0.09 | 0.17 | 0.14 | 0.08 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Notes: Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level. We also control for age, social capital , rain, soil suitability, topography, and distance to the capital. ACME stands for average causal mediation effects

## 6. Discussion and Conclusion

We find the incomes of Ghana's pineapple farmers not only to be explained by current factors but also by the experiences of their ancestors. That history matters in general has been found before (Acemoglu and Robinson 2001, Nunn 2008) and that it matters for Ghana's pineapple farmers is also not a new finding (Wuepper and Drosten 2016, Wuepper and Sauer 2016). The contribution of our analysis is to increase our understanding of the factors that enable the farmers to do better than predicted from the experiences of their ancestors. If ancestral experiences would fully determine incomes – e.g. through institutions or culture – it would not be possible to find such factors. However, we find that several farmers achieve higher incomes than we would predict from historic variables. These farmers are empowered by social and human capital, which have strong and robust effects, independent from which historical experience we analyze.

Using mediation analysis, we find that between 20 and 41% of the social capital effect runs through a more intensive production but less than 10% of the education effect is mediated by this channel. The remaining social capital effect could be better access to information, more reliable value chains, and/or improved access to (informal and formal) credit and insurance. The main channel for education could be improved applications of the same inputs, better understanding, different attitudes or better marketing.

There is widespread consensus that both human and social capital are important determinants of economic growth (Mankiw et al. 1992, Benhabib and Spiegel 1994, Knack and Keefer 1997, La Porta et al. 1997, Zak and Knack 2001, Knack and Zak 2003, Gennaioli et al. 2013, Wantchekon et al. 2015). However, this is the first empirical evidence that these are the factors that mitigate historically caused constrains to economic development. This complements the finding of Nunn (2013) that institutions can achieve the same. Notably, Michalopoulos and Papaioannou (2012) find that in Africa, the influence of national institutions quickly vanishes with the distance to the country's capital. As we are the first study to analyze the mitigation of historic constraints in a rural setting, it could well be that human and social capital are the main factors that support individuals in rural settings to overcome historic constraints. The comparison of such effects in urban and rural areas could be a topic for future research.

Future research could also clarify how human and social capital and institutions are connected to each other. First steps have already been undertaken – not in the context of overcoming historically caused constraints, but in the general context of economic development. Zak and Knack (2001) and La Porta et al. (1997) find that human capital increases social capital and Papagapitos and Riley (2009) find that social capital increases human capital. Obikili (2015) also finds this in colonial Western Africa. Acemoglu et al. (2014) argue that it is institutions that increase human capital, but they also argue that institutions have historic origins, referring to the work of Acemoglu and Robinson (2001) and Acemoglu et al. (2002). Furthermore, Michalopoulos and Papaioannou (2012) present evidence that in Sub-Saharan Africa, the influence of institutions drastically decreases with distance to the capital, so that residents of rural areas are not much influenced.

Thus, we might conclude that first of all, historically caused constraints to economic development can be overcome in different ways but institutions are an unlikely channel in rural settings. Even though the incomes of the pineapple farmers in Ghana are partially explainable with the historic experiences of their ancestors, either human or social capital are sufficient to achieve a higher than predicted income. Because both factors are well known and often targeted by policy, we can expect that historically caused constrains will vanish in time. Secondly, because human and social capital positively affect each other, policies that target one can achieve positive feedback effects on the other. Thirdly, at a national level, improving institutions can be expected to positively influence human and social capital, but improving human and social capital can also be expected to change a nation's instutions. Overall, despite finding historical persistence everywhere (Nunn 2013, Spolaore and Wacziarg 2013), our findings make hope that no spectacular effort and no innovative policy is necessary to mitigate historic constraints. Investing into human and social capital is a promising approach to speed up this process.

## APPENDIX

Table A3

| | (1) | (2) |
|---|---|---|
| | income | income |
| cereals | 0.0646** | |
| | (0.0256) | |
| slavery | | -0.306*** |
| | | (0.0221) |
| farming skills | 0.00589 | -0.000348 |
| | (0.0264) | (0.0210) |
| land | 0.106*** | 0.108*** |
| | (0.0389) | (0.0397) |
| loan | 0.0691** | 0.0667** |
| | (0.0307) | (0.0324) |
| price | 0.00386 | 0.00298 |
| | (0.0119) | (0.0101) |
| age | 0.00166 | -0.000586 |
| | (0.0248) | (0.0211) |
| edu | 0.0439*** | 0.0464*** |
| | (0.0154) | (0.0153) |
| rain | 0.0305*** | 0.0248*** |
| | (0.0105) | (0.00876) |
| soil | -0.00814 | -0.00330 |
| | (0.0101) | (0.00874) |
| elevation | -0.0113 | -0.00282 |
| | (0.0803) | (0.0638) |
| ruggedness | -0.0241 | -0.100* |
| | (0.0760) | (0.0539) |
| dist accra | -0.0331*** | -0.0854** |
| | (0.0125) | (0.0356) |
| R-sq | 0.29 | 0.30 |
| N | 398 | 398 |

Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district and ethnicity level.

Table A4

| | (1) | (2) |
|---|---|---|
| | income | income |
| cereals | 0.292** | |
| | (0.130) | |
| slavery | | -0.208* |
| | | (0.118) |
| rain | 0.0316* | 0.0379** |
| | (0.0185) | (0.0163) |
| soil | -0.0172 | -0.0182 |
| | (0.0223) | (0.0216) |
| elevation | -0.0202 | -0.0153 |
| | (0.0620) | (0.0586) |
| ruggedness | -0.0432 | -0.0366 |
| | (0.0542) | (0.0492) |
| dist accra | -0.167** | -0.124* |
| | (0.0798) | (0.0739) |
| R-sq | 0.104 | 0.157 |
| N | 398 | 398 |

Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district and ethnicity level.

Table A5

| | (1) | (2) |
|---|---|---|
| | Difference1 | Difference2 |
| export | 0.0630** | 0.0643** |
| | (0.0258) | (0.0284) |
| variety | 0.0946*** | 0.0856** |
| | (0.0240) | (0.0368) |
| residues | 0.0701*** | 0.0681*** |
| | (0.0110) | (0.0123) |
| mulching | 0.0322*** | 0.0327*** |
| | (0.0105) | (0.0114) |
| fertilizer | 0.0898** | 0.0825** |
| | (0.0395) | (0.0407) |
| price | -0.0147 | -0.0143 |
| | (0.0164) | (0.0161) |
| skills | 0.0103 | 0.0127 |
| | (0.0219) | (0.0244) |
| age | -0.000919 | -0.000893 |
| | (0.000919) | (0.000876) |
| education | 0.196 | -0.116 |
| | (0.190) | (0.158) |
| rain | -0.0171* | -0.0164* |
| | (0.00913) | (0.00952) |
| soil | 0.0431*** | 0.0291* |
| | (0.0144) | (0.0170) |
| elevation | -0.0249 | -0.00463 |
| | (0.0184) | (0.0140) |
| ruggedness | 0.0602 | -0.0470 |
| | (0.0834) | (0.0909) |
| dist accra | -0.214*** | -0.214*** |
| | (0.0258) | (0.0230) |
| R-sq | 0.36 | 0.43 |
| N | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A6

| | (1) | (2) |
|---|---|---|
| | Difference1 | Difference2 |
| leader inno | 0.0462** | 0.0461** |
| | (0.0192) | (0.0209) |
| leader trad | -0.0415** | -0.0386** |
| | (0.0188) | (0.0191) |
| edu | 0.0595*** | 0.0576*** |
| | (0.0107) | (0.00992) |
| social capital | 0.0634* | 0.0643* |
| | (0.0339) | (0.0336) |
| price | -0.00848 | -0.00824 |
| | (0.00784) | (0.00847) |
| skills | -0.00601 | -0.00334 |
| | (0.0347) | (0.0427) |
| age | -0.000584 | -0.000630 |
| | (0.00107) | (0.001000) |
| rain | -0.0420*** | -0.0410*** |
| | (0.0122) | (0.0106) |
| soil | 0.0428*** | 0.0281*** |
| | (0.00936) | (0.00911) |
| elevation | 0.00366 | 0.0171 |
| | (0.0213) | (0.0286) |
| ruggedness | -0.0121 | -0.116 |
| | (0.0919) | (0.100) |
| dist accra | -0.174*** | -0.188*** |
| | (0.0443) | (0.0456) |
| R-sq | 0.32 | 0.40 |
| N | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A7

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Difference1 | Difference1 | Difference2 | Difference2 |
| export | 0.0443 | | 0.104 | |
| | (0.0604) | | (0.0726) | |
| edu | 0.0571*** | 0.0560*** | 0.0472*** | 0.0531*** |
| | (0.0131) | (0.000709) | (0.0109) | (0.00374) |
| social capital | | 0.115*** | | 0.115*** |
| | | (0.0201) | | (0.0187) |
| skills | 0.0182* | -0.00643 | 0.0158 | -0.00668 |
| | (0.0104) | (0.0251) | (0.0135) | (0.0282) |
| age | 0.000368 | -0.0000919 | 0.000548 | 0.000169 |
| | (0.000235) | (0.000201) | (0.000355) | (0.000400) |
| price | 0.335*** | 0.364*** | 0.265* | 0.339** |
| | (0.112) | (0.118) | (0.154) | (0.153) |
| rain | -0.0381*** | -0.0530*** | -0.0307*** | -0.0466*** |
| | (0.00277) | (0.00370) | (0.00432) | (0.00433) |
| soil | 0.0409*** | 0.0466*** | 0.0405** | 0.0432*** |
| | (0.00766) | (0.00145) | (0.0184) | (0.0119) |
| elevation | 0.0154 | 0.0132 | -0.00931 | -0.0120 |
| | (0.00957) | (0.0259) | (0.0440) | (0.0608) |
| ruggedness | -0.0128 | -0.0539 | -0.0791 | -0.123 |
| | (0.0529) | (0.0737) | (0.0827) | (0.104) |
| dist accra | -0.107*** | -0.193*** | -0.112*** | -0.224*** |
| | (0.00989) | (0.0445) | (0.0264) | (0.0650) |
| R-sq | 0.30 | 0.30 | 0.31 | 0.31 |
| N | 398 | 398 | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A8

| | (1) | (2) |
|---|---|---|
| | Difference1 | Difference2 |
| | | |
| Leader 1 | 0.0888 | 0.0827 |
| | (0.0713) | (0.0718) |
| leader | | |
| | | |
| edu | 0.0652*** | 0.0631*** |
| | (0.0163) | (0.0157) |
| price | -0.00640 | -0.00562 |
| | (0.0143) | (0.0143) |
| skills | -0.00131 | 0.00307 |
| | (0.0262) | (0.0268) |
| age | -0.00411 | -0.00414 |
| | (0.0165) | (0.0165) |
| rain | -0.0337* | -0.0328* |
| | (0.0185) | (0.0185) |
| soil | 0.0451*** | 0.0295* |
| | (0.0162) | (0.0162) |
| elevation | 0.0105 | 0.0237 |
| | (0.0506) | (0.0516) |
| ruggedness | 0.00548 | -0.0981 |
| | (0.0778) | (0.0784) |
| dist accra | -0.120 | -0.134 |
| | (0.161) | (0.162) |
| R-sq | 0.288 | 0.374 |
| N | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A9

| | (1) | (3) | (2) | (4) |
|---|---|---|---|---|
| | Difference1 | Difference1 | Difference2 | Difference2 |
| Social capital | 0.0982*** | 0.109*** | 0.0971*** | 0.108*** |
| | (0.0264) | (0.0284) | (0.0267) | (0.0288) |
| leader | 0.0129 | 0.0252 | 0.00794 | 0.0204 |
| | (0.0200) | (0.0244) | (0.0200) | (0.0246) |
| Nonfarm inc. | 0.175*** | | 0.176*** | |
| | (0.0501) | | (0.0498) | |
| land | 0.0835*** | | 0.0854*** | |
| | (0.0240) | | (0.0245) | |
| export | 0.0293 | | 0.0306 | |
| | (0.0340) | | (0.0342) | |
| loan | 0.0887*** | | 0.0935*** | |
| | (0.0256) | | (0.0256) | |
| education | 0.0332** | 0.0586*** | 0.0304** | 0.0566*** |
| | (0.0160) | (0.0161) | (0.0154) | (0.0156) |
| age | -0.00849 | -0.00441 | -0.00876 | -0.00455 |
| | (0.0164) | (0.0185) | (0.0161) | (0.0185) |
| price | -0.0168 | -0.0141 | -0.0163 | -0.0134 |
| | (0.0119) | (0.0132) | (0.0120) | (0.0134) |
| skills | -0.0118 | -0.00843 | -0.00773 | -0.00431 |
| | (0.0192) | (0.0189) | (0.0195) | (0.0194) |
| rain | -0.0609*** | -0.0519*** | -0.0599*** | -0.0508*** |
| | (0.0179) | (0.0180) | (0.0179) | (0.0181) |
| soil | 0.0417*** | 0.0463*** | 0.0260* | 0.0308* |
| | (0.0155) | (0.0175) | (0.0153) | (0.0176) |
| elevation | 0.0482 | 0.00584 | 0.0628 | 0.0191 |
| | (0.0561) | (0.0513) | (0.0570) | (0.0526) |
| ruggedness | -0.0420 | -0.0315 | -0.146* | -0.135 |
| | (0.0855) | (0.0885) | (0.0853) | (0.0888) |
| dist accra | -0.277 | -0.213 | -0.294 | -0.226 |
| | (0.185) | (0.175) | (0.186) | (0.175) |
| R-sq | 0.427 | 0.304 | 0.499 | 0.385 |
| N | 398 | 398 | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A10

| | (1) | (2) |
|---|---|---|
| | Difference1 | Difference2 |
| social capital | 0.112*** | 0.111*** |
| | (0.0306) | (0.0331) |
| education | 0.0579*** | 0.0560*** |
| | (0.0146) | (0.0172) |
| price | -0.0133 | -0.0126 |
| | (0.0128) | (0.0121) |
| ability | -0.00224 | 0.000780 |
| | (0.0256) | (0.0251) |
| age | -0.00210 | -0.00265 |
| | (0.0165) | (0.0176) |
| rain | -0.0541*** | -0.0526*** |
| | (0.0193) | (0.0193) |
| soil | 0.0438*** | 0.0288*** |
| | (0.00933) | (0.00869) |
| elevation | 0.00611 | 0.0193 |
| | (0.0468) | (0.0462) |
| ruggedness | -0.0350 | -0.138*** |
| | (0.0367) | (0.0371) |
| dist accra | -0.222 | -0.234 |
| | (0.151) | (0.156) |
| R-sq | 0.295 | 0.379 |
| N | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A11

| | (1) | (2) |
|---|---|---|
| Sample | Without Kwahu | |
| | Difference1 | Difference2 |
| social capital | 0.120*** | 0.118*** |
| | (0.0311) | (0.0311) |
| education | 0.0635*** | 0.0618*** |
| | (0.0162) | (0.0157) |
| skills | -0.00918 | -0.00572 |
| | (0.0207) | (0.0212) |
| age | 0.000181 | 0.000124 |
| | (0.00204) | (0.00203) |
| price | 0.0447*** | 0.0426** |
| | (0.0168) | (0.0171) |
| rainfall | -0.0532*** | -0.0515*** |
| | (0.0197) | (0.0198) |
| soil | 0.0475*** | 0.0324* |
| | (0.0174) | (0.0174) |
| elevation | 0.0127 | 0.0254 |
| | (0.0511) | (0.0525) |
| ruggedness | -0.0576 | -0.159* |
| | (0.0909) | (0.0909) |
| dist accra | -0.201 | -0.214 |
| | (0.164) | (0.165) |
| R-sq | 0.30 | 0.38 |
| N | 371 | 371 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A12

| | (1) |
|---|---|
| | income |
| Sample | Without Cereal Regions |
| suitability cereals | -0.0621 |
| | (0.0400) |
| land | 0.0952** |
| | (0.0375) |
| loan | 0.0731* |
| | (0.0381) |
| education | 0.0473** |
| | (0.0216) |
| skills | -0.0212 |
| | (0.0361) |
| price | -0.00243 |
| | (0.0140) |
| age | 0.00430 |
| | (0.0201) |
| rain | 0.0258*** |
| | (0.00549) |
| soil | 0.0118 |
| | (0.0142) |
| elevation | -0.192 |
| | (0.139) |
| ruggedness | 0.167*** |
| | (0.0411) |
| dist accra | -0.137*** |
| | (0.0463) |
| R-sq | 0.335 |
| N | 272 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A13

| | (1) | (2) |
|---|---|---|
| | Difference1 | Difference2 |
| slavery | -0.158 | |
| | (0.104) | |
| cereals | | 0.0711 |
| | | (0.101) |
| price | 0.0686*** | 0.0681*** |
| | (0.0249) | (0.0246) |
| skills | 0.0147 | 0.0156 |
| | (0.0428) | (0.0430) |
| age | -0.000235 | -0.000240 |
| | (0.00204) | (0.00201) |
| education | 0.0904 | -0.0248 |
| | (0.167) | (0.287) |
| rain | -0.0240 | -0.0241 |
| | (0.0162) | (0.0162) |
| soil | 0.0353* | 0.0358 |
| | (0.0213) | (0.0219) |
| elevation | 0.0175 | 0.0174 |
| | (0.0699) | (0.0711) |
| ruggedness | -0.00847 | -0.00663 |
| | (0.0739) | (0.0727) |
| dist accra | -0.0767 | -0.0765 |
| | (0.176) | (0.176) |
| R-sq | 0.268 | 0.269 |
| N | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A14

| | (1) |
|---|---|
| | income |
| government schools | 0.0105 |
| | (0.0214) |
| railroad distance | 0.0339 |
| | (0.0891) |
| rain | 0.0365*** |
| | (0.00950) |
| soil | -0.0186 |
| | (0.0118) |
| elevation | -0.0144 |
| | (0.0689) |
| ruggedness | -0.0435 |
| | (0.0698) |
| distance accra | -0.120*** |
| | (0.0150) |
| coast distance | 0.110*** |
| | (0.0311) |
| R-sq | 0.157 |
| N | 398 |

Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district and ethnicity level.

Table A15

| | (1) | (2) |
|---|---|---|
| | income | income |
| tenure | 0.0149 | |
| | (0.0114) | |
| rented land | | 0.0251 |
| | | (0.0478) |
| rain | 0.0448*** | 0.0446*** |
| | (0.0105) | (0.0109) |
| soil | -0.0208* | -0.0239*** |
| | (0.0109) | (0.00890) |
| elevation | -0.0253 | -0.0247 |
| | (0.0830) | (0.0856) |
| ruggedness | 0.0263 | 0.0312 |
| | (0.0807) | (0.0749) |
| dist accra | -0.0722*** | -0.0748*** |
| | (0.0142) | (0.00524) |
| R-sq | 0.154 | 0.153 |
| N | 398 | 398 |

Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district and ethnicity level.

Table A16

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Difference 1 | Difference 1 | Difference 1 | Difference 2 | Difference 2 | Difference 2 |
| social capital | 0.0300* | 0.0310* | 0.0404** | 0.0387** | 0.0402** | 0.0507*** |
| | (0.0156) | (0.0186) | (0.0170) | (0.0163) | (0.0201) | (0.0181) |
| intensification | 0.294*** | | | 0.305*** | | |
| | (0.0663) | | | (0.0774) | | |
| fertilizer | | 0.326*** | | | 0.330*** | |
| | | (0.0629) | | | (0.0782) | |
| export | | | 0.0919*** | | | 0.0847** |
| | | | (0.0289) | | | (0.0325) |
| edu | 0.0505*** | 0.0565*** | 0.0507*** | 0.0536*** | 0.0597*** | 0.0543*** |
| | (0.0180) | (0.0155) | (0.0182) | (0.0190) | (0.0162) | (0.0199) |
| age | 0.00165 | 0.00771 | 0.00263 | 0.00508 | 0.0114 | 0.00691 |
| | (0.0182) | (0.0199) | (0.0208) | (0.0206) | (0.0231) | (0.0232) |
| rain | -0.0203 | -0.0133 | -0.0208 | -0.0238 | -0.0165 | -0.0236 |
| | (0.0169) | (0.0160) | (0.0176) | (0.0184) | (0.0173) | (0.0195) |
| soil | 0.0299 | 0.0271 | 0.0228 | 0.0335* | 0.0300 | 0.0243 |
| | (0.0181) | (0.0173) | (0.0178) | (0.0196) | (0.0187) | (0.0189) |
| elevation | -0.0137 | -0.0262 | -0.00856 | -0.0128 | -0.0255 | -0.00818 |
| | (0.0570) | (0.0600) | (0.0580) | (0.0634) | (0.0684) | (0.0652) |
| ruggedness | 0.0332 | -0.0476 | 0.0142 | 0.0108 | -0.0728 | -0.0149 |
| | (0.0802) | (0.0705) | (0.0761) | (0.0971) | (0.0774) | (0.0912) |
| dist accra | 0.0985 | 0.0378 | 0.0738 | 0.0910 | 0.0278 | 0.0599 |
| | (0.0820) | (0.0729) | (0.0781) | (0.0771) | (0.0701) | (0.0713) |
| R-sq | 0.170 | 0.150 | 0.092 | 0.167 | 0.143 | 0.083 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

Table A17

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Difference 1 | Difference 1 | Difference 1 | Difference 2 | Difference 2 | Difference 2 |
| edu | 0.0505*** | 0.0565*** | 0.0507*** | 0.0536*** | 0.0597*** | 0.0543*** |
| | (0.0180) | (0.0155) | (0.0182) | (0.0190) | (0.0162) | (0.0199) |
| intensification | 0.294*** | | | 0.305*** | | |
| | (0.0663) | | | (0.0774) | | |
| fertilizer | | 0.326*** | | | 0.330*** | |
| | | (0.0629) | | | (0.0782) | |
| export | | | 0.0919*** | | | 0.0847** |
| | | | (0.0289) | | | (0.0325) |
| social capital | 0.0300* | 0.0310* | 0.0404** | 0.0387** | 0.0402** | 0.0507*** |
| | (0.0156) | (0.0186) | (0.0170) | (0.0163) | (0.0201) | (0.0181) |
| age | 0.00165 | 0.00771 | 0.00263 | 0.00508 | 0.0114 | 0.00691 |
| | (0.0182) | (0.0199) | (0.0208) | (0.0206) | (0.0231) | (0.0232) |
| rain | -0.0203 | -0.0133 | -0.0208 | -0.0238 | -0.0165 | -0.0236 |
| | (0.0169) | (0.0160) | (0.0176) | (0.0184) | (0.0173) | (0.0195) |
| soil | 0.0299 | 0.0271 | 0.0228 | 0.0335* | 0.0300 | 0.0243 |
| | (0.0181) | (0.0173) | (0.0178) | (0.0196) | (0.0187) | (0.0189) |
| elevation | -0.0137 | -0.0262 | -0.00856 | -0.0128 | -0.0255 | -0.00818 |
| | (0.0570) | (0.0600) | (0.0580) | (0.0634) | (0.0684) | (0.0652) |
| ruggedness | 0.0332 | -0.0476 | 0.0142 | 0.0108 | -0.0728 | -0.0149 |
| | (0.0802) | (0.0705) | (0.0761) | (0.0971) | (0.0774) | (0.0912) |
| dist accra | 0.0985 | 0.0378 | 0.0738 | 0.0910 | 0.0278 | 0.0599 |
| | (0.0820) | (0.0729) | (0.0781) | (0.0771) | (0.0701) | (0.0713) |
| R-sq | 0.170 | 0.150 | 0.092 | 0.167 | 0.143 | 0.083 |
| N | 398 | 398 | 398 | 398 | 398 | 398 |

Difference1 is the difference between the actual income of the farmers and their predicted income from the historic impact of the transatlantic slave trade. Difference2 is the difference between the actual income of the farmers and their predicted income from which was historically the main subsistence crop of their ancestors. Standard errors are multi-dimensionally clustered at the ethnicity and farm group level, fixed effects control for unobserved influences at the district, farm group, and ethnicity level.

# Chapter 5
# The Diffusion of Sustainable Intensification Practices in Ghana:
## Why is Mulching so much more Common than the use of Organic Fertilizers?

with Johannes Sauer and Linda Kleemann[6]

**Abstract** Sustainable intensification is highly profitable for Ghana's pineapple farmers - especially when several technologies are combined. Because such technologies are generally knowledge intensive, they are widely promoted through trainings. However, whereas mulching is diffused amongst 50% of the farmers, the use of sustainable crop rotations, cover crops, or manures is only diffused amongst 15% of the farmers. Using data from a representative sample of 400 farmers, we find that it is the characteristics of the technologies explain this pattern. Mulching is one of the simplest sustainable intensification technologies, having similar knowledge requirements as conventional practices. It thus diffuses relatively easily through peer learning. The other practices require more individual adaptations, which makes peer learning less effective. This explains why the technologies are so differently common and why we find that training only has a significant effect on the adoption of organic fertilizers but not on mulching (anymore).

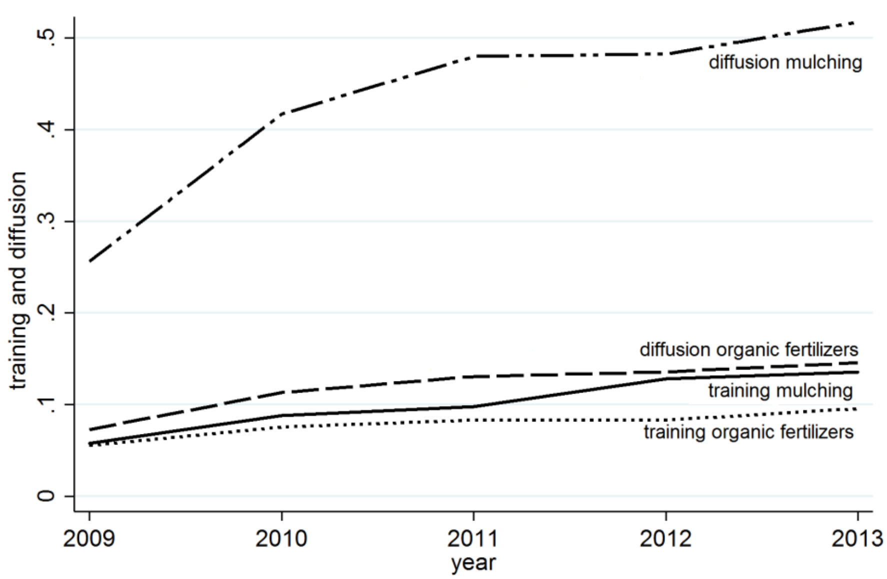
**Contributions** David Wuepper collected the data, performed the analysis and wrote the article. Linda Kleemann greatly supported the data collection in Ghana and contributed her own data from 2010. Linda Kleemann and Johannes Sauer also contributed to the article with ideas, feedback, and discussions. An early version of the chapter was presented at the Bioecon conference at the University of Cambridge, where valuable feedback was given.

[6]The Kiel Institute for the World Economy, Group Poverty Reduction, Equity, and Development

## 1. Introduction

Sustainable intensification can be defined as "producing more output from the same area of land while reducing the negative environmental impacts and at the same time increasing contributions to natural capital and the flow of environmental services" (Pretty et al. 2011). African agriculture has a high potential for sustainable intensification (McIntyre et al. 2009, Pretty et al. 2011, Tilman et al. 2011) and pineapple farming in Ghana is no exception (Kleemann and Abdulai 2013). Because the economic costs of extensive production and land degradation are high in Ghana (Diao and Sarpong 2007, World Bank 2011) sustainable intensification is actively promoted by extension services, processing companies, international NGOs and development organizations (German Society for International Cooperation 2005, USAID 2009, Government of Ghana 2010, Millenium Development Authority 2011, McMillan 2012, USAID 2013). Nevertheless, the adoption rate is rather low - with the exception of mulching, which is used by about half of Ghana's pineapple farmers by now (see figure 1).

**Figure 1. Diffusion of Organic Practices**




Notes: Data from own survey, based on sampled pineapple farmers in the South of Ghana

Mulching can mean a spectrum of techniques that all have in common that bare soils are avoided by covering them up with any material to suppress weed growth and conserve soil moisture (Awodoyin et al. 2007, Dzomeku et al. 2009, Snapp and Pound 2011). Used materials might be plastic foil, organic materials, or textiles. Important is that mulching material is not living and used for the sole purpose of covering bare soils.

Other available sustainable intensification technologies in Ghana include crop rotations that include legumes and other crops that increase the nutrient content of the soil, to intercrop those plants together with the main crop, and the use of crop residues and other organic materials as natural fertilizers. Those practices might be called organic fertilizers, as their main purpose is to enrich the nutrient content of the soil. As can be seen in figure 1, these practices are far less diffused than mulching even though both have been shown to be similarly profitable (Kleemann and Abdulai 2013). It has also been found that combining mulching and organic fertilizers non-linearly increases profits (Kleemann and Abdulai 2013), which suggests that a wider diffusion of especially organic fertilizers could pay off greatly.

The large literature on the diffusion of agricultural innovations suggests several explanations why seemingly profitable innovations do not quickly diffuse amongst the farmers and why the diffusion of some innovations is slower than that of others (Feder et al. 1985, Anderson and Feder 2004, Foster and Rosenzweig 2010). Such explanations include heterogeneous profits, such that not all farmer actually benefit from adoption (Suri 2011), uninsured risk (Dercon and Christiaensen 2011, Karlan et al. 2012), insecure tenure rights (Abdulai et al. 2011, Fenske 2011), and bounded rationality (Duflo et al. 2011, Wuepper et al. 2016). An especially prominent explanation are information disequilibria. Farmers need to learn about the existence, profitability, and correct application of new technologies, before they are able and willing to adopt them. Thus, if not all farmers have access to the same amount of information, it is suggested that for profitable innovations, the farmers with better information access adopt first, and the others only follow when they have received sufficient information themselves (Bandiera and Rasul 2006, Conley and Udry 2010, Kabunga et al. 2012). The main information sources in developing countries are usually other farmers and trainings (Moser and Barrett 2006, Dercon et al. 2009, Rogers 2010, Pan et al. 2015).

In the context of this study, we investigate whether the nature of the different sustainable intensification technologies explains why some are more commonly used than others. Specifically, does the complexity organic fertilizers prevent a diffusion similar to that of mulching?

As a first indicator, the farmers describe mulching to be much easier to learn from others than the use of organic fertilizers. The reason is that mulching requires far less knowledge and individual adaptation compared to the use of organic fertilizers. If a farmer observes that a neighbor profitably uses mulching, she can imitate this neighbor and also use mulching. Even if the neighbor uses a material (say plastic foil) that is not available to the farmer, she can simply use a different material (say grass or straw). In contrast, learning that a neighbor profitably integrates legumes in her crop rotation requires the potential adopter to learn about all the requirements of that legume and how that crop interacts with the currently used farming practices, e.g. the use of agro-chemicals. Furthermore, as many organic fertilizers are living plants, they might do differently well on different plots as a function of soil moisture, micro-climate, disease pressure and soil nutrients (Snapp and Pound 2011). Investigating the implications of such differences amongst sustainable intensification technologies is a contribution to two strains of literature. First, the literature on sustainable intensification commonly emphasizes the increased knowledge intensity of the involved technologies *compared to conventional farming systems*, whereas we are the first to focus on knowledge differences *amongst* sustainable intensification technologies. Secondly, the literature on the diffusion of innovations has recently been augmented by the finding that training is effective to start the innovation diffusion process, but not to drive it at later stages (Krishnan and Patnam 2014). We augment the literature by showing that the effectiveness of training also depends on the characteristic of the technology. The more complex an innovation, the less effective is peer learning and the more important becomes training. Thus, the optimal amount of provided training is a function of how difficult it is for farmers to learn from their peers.

In our case, we estimate that to diffuse the use of organic fertilizers as much as mulching currently is, trainings would need to increase from currently 10% trained farmers to about 25% trained farmers (to achieve 50% diffusion).

As we do not conduct a randomized control trial, we must carefully consider the issue of unobserved heterogeneity. We use 2SLS to control for the endogeneity of estimated peer effects and selection into trainings by relying on an approach similar to the one suggested by Bramoullé et al. (2009), which has also been used by Krishnan and Patnam (2014), and we also incorporate ideas from Zeitlin (2011), Munshi (2004), and Caeyers and Fafchamps (2016).

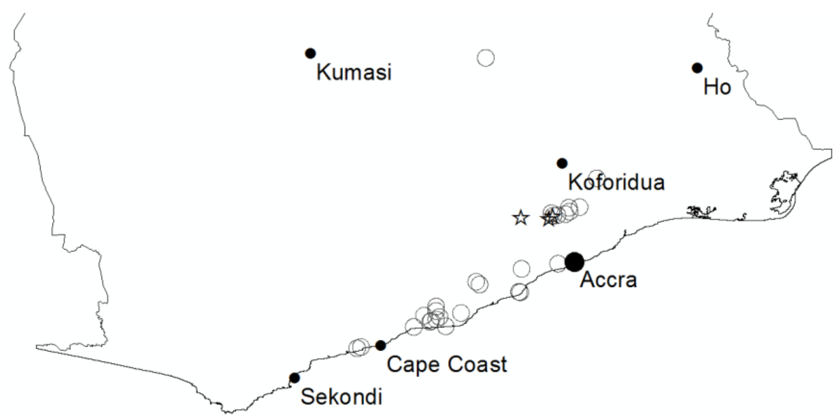
In the next section (2), we provide some background information and describe our data. We then explain our empirical framework in section 3 and present our results in section 4. In section 5, we discuss and conclude the study.

## 2. Context and Data

The pineapple farmers of Ghana have received a lot of academic attention in recent years (Udry and Conley 2004, Udry and Anagol 2006, Goldstein and Udry 2008, Conley and Udry 2010, Suzuki et al. 2011, Gatune et al. 2013, Kleemann and Abdulai 2013, Wuepper et al. 2016). One reason is the dynamism of the sector, starting with the business opportunity to grow pineapple for export to the European Union in the 1990s. Most of the pineapple farmers had previously relied on far less profitable roots, tubers, and cereals –with large shares for own consumption. With the decision to grow pineapples for export, the farmers needed to learn about how to intensify their former extensive production systems (Conley and Udry 2010). They also needed to learn about new business arrangements and the differences between formal and informal contracts (Suzuki et al. 2011, Gatune et al. 2013, Wuepper and Sauer 2016). The performance of Ghana's pineapple sector was quite strong until demand in the European Union switched around the year 2004 to a new variety, which is more expensive, environmentally sensitive and requires more inputs (Fold and Gough 2008). Many formerly successful companies went out of business, and many farmers made critical losses (Barrett et al. 2012). Currently, the remaining pineapple farmers and processing companies face the double challenge to intensify pineapple production, in order to be profitable, while avoiding negative side-effects, especially from reduced environmental services and adverse health effects (Kleemann and Abdulai 2013). The incentive for the sustainable intensification of the pineapple production is the growing organic market worldwide (Kleemann et al. 2014) and the growing domestic market, including the rise of supermarkets (World Bank 2011). The farmers are supported by a range of national and especially international stakeholders. The US-American USAID, the German GIZ, the World Bank, The United Nations Millenium Development Authority, various NGOs, some of the processing companies, and Ghana's ministry of food and agriculture (MOFA), all provide training to the farmers, in such diverse topics as farm management, accounting, input use, and farming practices. Trainings are often provided together by an international stakeholder and Ghana's MOFA. Similar to Ghana's cocoa growers, the pineapple farmers are almost all organized in farmers' groups, to support each other and to obtain access to credit, trainings, and other external production inputs. These farmers' groups are locally organized, so that their members can frequently meet. However, their sizes vary, so that sometimes, farmers from several communities are organized within one farmers' group, and sometimes, there are multiple farmers' groups in one community. For a lack of information about the communities, trainings can usually not be strictly targeted, e.g. towards farmers with more need, potential, or interest, but the strategy is to start training farmers in one community and then to move on to the next community. Most communities receive several trainings a year from various stakeholders. In addition to the trainings provided by development organizations, processing companies such as especially Blue Skies provide training specifically for the communities where they have suppliers, whereas many NGOs specifically target more remote communities.

Acknowledging the potential of sustainable intensification practices, many trainings focus on mulching (to conserve soil moisture) and a range of organic fertilizers, such as incorporating certain leguminous crops into the crop rotation, to interplant such crops together with pineapple, to use crop residues, or animal manure. Increasing soil nutrient contents and moisture increase fruit quality and quantity (Norman 1986) and specifically for the farmers surveyed for this study, Kleemann and Abdulai (2013) find that sustainable intensification practices are highly profitable.

**Figure 2. Sampling**




To investigate the effect of trainings and peer-learning on the diffusion of such practices, we surveyed 400 farmers in 2013. Half of these farmers where surveyed already in 2010 by Kleemann and Abdulai (2013), the other half was interviewed the first time. The farmers from the first period are farmers who where certified for export at that time. As lists of such farmers are readily available, a three stage stratified sampling procedure was feasible, starting with the districts were most pineapples are produced (in the Eastern Region, the Central Region, and Greater Accra), followed by the farmers' groups that are certified to export pineapples, and finishing with a proportional sampling of individual farmers according to the number of pineapple producers in each group. For non-certified pineapple farmers, there are no reliable statistics available, so the sampling is based on the information provided by development agencies and extension agents. When selecting non-certified farmers without lists, special emphasis was placed on the representativeness of the farmers, so as not to disproportionally sample "easier to reach" farmers.

Nevertheless, very remote farmers are likely to be underrepresented because they are less likely to be on a list for the stratified random sampling and for the other farmers, the probability to be sampled decreased with the distance to those farmers. However, such farmers are also less likely to grow pineapple in any relevant quantity, so we judge the final sample as representative for the pineapple farmers of Ghana (which, as can be seen in figure 2 are concentrated near the coast, as pineapple growing conditions deteriorate quickly north of Kumasi). The final sample size is 398 farmers, of whom roughly half have been export certified at some point in time and the other half was never certified. Figure 2 shows the south of Ghana. Circles indicate sampling locations. The stars indicate the locations of the main pineapple processors (three, very close together, below Koforidua), black dots show major cities.

**Table 1. Variables and Summary Statistics**

| variable | description | mean | sd | min | max |
|---|---|---|---|---|---|
| adoption organic fertilizer | Binary, whether or not an organic fertilizer is used on any field | .11 | .32 | 0 | 1 |
| adoption mulch | Binary, whether or not mulching is used on any field | .43 | .49 | 0 | 1 |
| training organic fertilizer | Binary, whether the farmer was trained in organic fertilizers until this period | .07 | .26 | 0 | 1 |
| training mulch | Binary, whether the farmer was trained in mulching until this period | .10 | .30 | 0 | 1 |
| rain | Reported rainfall quantity from 1 = problematic to 6 = optimal | 4.49 | 1.39 | 1 | 6 |
| soil | Reported soil fertility from 1 = no constraint to 4 = important constraint | 1.64 | .76 | 1 | 4 |
| age | Age of the farmer in years | 43.28 | 10.82 | 18 | 77 |
| edu | Education level of the farmer, from 1 = none to 6 = University, plus 7 = other | 2.70 | 1.19 | 1 | 7 |
| farmsize | Hectares potentially available to grow pineapple, including currently not used | 3.90 | 5.21 | .5 | 50 |
| risk pref | From a choice experiment, 1= most risk averse to 6 = least risk averse | 3.31 | 1.34 | 1 | 6 |
| nonfarm | Importance of nonfarm income, from 1= non-existent to 6 = important | 2.04 | 1.63 | 1 | 6 |
| loan | Binary, whether the farmer received a credit or not | .21 | .41 | 0 | 1 |
| contract | Binary, whether the farmer is in a formal contract arrangement | .18 | .39 | 0 | 1 |

In order to more precisely estimate the causal effect of training and peer learning, we asked the farmers for each of their plots whether and when they adopted a sustainable intensification practice and we asked them when they received training, from whom, and about what topic. We also asked about annual values for some of our control variables, such as rainfall, prices, credit, and contract farming. We consider the last five years, between 2009 and 2013, as recall error is likely to increase in time and because we have data for 2009 and 2013 for half of the farmers as a sanity check. Overall, we thus use 1990 observations, for 398 farmers and 5 periods. Table 1 presents the main variables and summary statistics. It can be seen that currently, 43% of the farmers use mulching but only 11% use organic fertilizers (of those 11%, 72% use leguminoses in their crop rotation, 51% use crop residues, 35% use intercropping, and 21% use other organic fertilizers). In contrast to the wide difference in adoption, training in mulching and training in organic fertilizers is provided to 10 and 7% of the farmers, respectively. The average farmer in our sample is 43 years old, male (women are almost entirely absent), has only completed Junior Secondary School, and does neither have much nonfarm income, nor a credit (only 21% do), or a contract arrangement with a company (only 18% do).

**Table 2. Information Sources**

| extension | neighbors | friends | group |
|---|---|---|---|
| 0,71 | 0,06 | 0,16 | 0,30 |

According to the farmers (table 2), they mostly discuss their farming practices with extension agents (71%), followed by members of their farmers' group (30%), their friends (16%), and their neighbors (6%)

## 3. Empirical Framework

It is well known that the identification of social interactions poses a range of identification challenges (Manski 2000, Moffitt 2001, Blume and Durlauf 2006), such as the reflection problem described by Manski (1993), and the problem of exclusion bias (Guryan et al. 2009, Caeyers and Fafchamps 2016). In this section, we will discuss the empirical challenges first for the estimation of peer-learning and then for the effect of training. Then we discuss our analytical framework and how we address the discussed challenges.

### i. The Identification of Peer-Learning Effects

In most economic research, social interactions are not directly observed, but only indirectly inferred from observed outcomes in the peer networks (Manski 2000). Identifying the right network is already the first challenge (Maertens and Barrett 2013). Foster and Rosenzweig (1995) and Munshi (2004) assume that peer-networks are the villages of the farmers. However, networks usually differ across contexts and e.g. Conley and Udry (2010) show that when it comes to learning about an innovation, networks are smaller than the villages. An expensive alternative is used by Van den Broeck and Dercon (2011), who take a full census of each village and ask the farmers about their contacts. Less expensive, Bandiera and Rasul (2006) ask the farmers to list a small number of their peers from whom they learn. In a comparable approach, Conley and Udry (2010) randomly match a small number of farmers and ask them about each others' behaviors and outcomes. Finally, Krishnan and Patnam (2014) use spatial proximity of their sampled farmers and define peer-networks to be within 1Km distance from each other. It is well known that misrepresenting the peer networks can bias the estimates of the peer effect (Maertens and Barrett 2013). In our context, the choice of the peer network is aided by the fact that the pineapples farmers are organized in local farmers' groups, which are a viable approximation of their peer network (see table 2).

As it is common to use the (often lagged) outcome of one's peers as opportunity for peer-learning, we directly run into the reflection problem described by Manski (1993) and the exclusion bias described by Guryan et al. (2009) and Caeyers and Fafchamps (2016). The former describes that homogeneous behavior within a peer network has more potential explanations than just peer learning. We need to distinguish between contextual effects (individuals in the same context tend to behave similar), endogenous effects (peer learning and other externalities), and correlated effects (individuals in the same peer group share common characteristics that also produce similar behaviors). The name reflection bias comes from the example that without additional information, it is impossible to know whether a mirror "reacts" to the person in front or the other way round, as both change simultaneously. Formally, we are interested in the following model:

$$y_{it} = \beta_1 x_{it} + \beta_2 \frac{\sum_{j \in \eta_i} y_{jt}}{\eta_i} + \beta_3 \frac{\sum_{j \in \eta_i} x_{jt}}{\eta_i} + u_{it} \quad (1)$$

$y_{it}$ denotes the outcome of individual $i$ at time $t$ (e.g. the adoption of an innovation), $x_{it}$ denotes her characteristics (e.g. age and education), $\frac{\sum_{j \in \eta_i} y_{jt}}{\eta_i}$ is the average outcome in her network $\eta_i$ (e.g. the average adoption rate of the innovation, excluding farmer $i's$ choice), and $\frac{\sum_{j \in \eta_i} x_{jt}}{\eta_i}$ are the average characteristics of her network without herself (e.g. the average farm size or amount of rainfall). Thus, we are interested in precisely identifying $\beta_2$, which is the causal effect of learning from peers on the adoption probability of the innovation. As we will further elaborate on below, we loosely follow the approach of Bramoullé et al. (2009) and use the lagged treatment and outcome of indirect neighbors to exogenize $\beta_2$.

An issue that has not yet received much attention is the exclusion bias that is created when OLS is used to estimate $\beta_2$. The exclusion bias is created when each farmer is excluded from the calculation of her peer statistics (the intuition is that a individuals cannot be their own peers, so they are excluded from the calculation of peer outcomes). This creates a systematic, negative correlation between the characteristics of the peers and the characteristics of the individual, which biases the OLS estimated peer effects downwards (Guryan et al. 2009) but which does not affect specifications that use the lagged outcome of the peers while controlling for the farmer's own lagged outcome and specific other set ups (Caeyers and Fafchamps 2016). In addition to alternative to substantive explanations such as negative assortative matching in the endogenous peer group formation, the exclusion bias can explain why OLS estimates of peer learning are usually considerably smaller than their corresponding instrumental variables estimates (Zeitlin 2011, Krishnan and Patnam 2014). In our case, we use the lagged outcome of the peers while controlling for own lagged outcome.

### ii. The Identification of Training Effects

The identification of training effects would be easier if we could assume that trainings are received fully randomly, which is however somewhat implausible, even if trainings are not precisely targeted. There is a likely degree of two-way selection. First, even if most trainers do not know what exact training is best in a certain community, some might do. Furthermore, even if farmers are generally highly interested in all trainings and thus, participation in a given locality is close to 100%, the incentive to join a training is higher for more interested farmers. Thus, as argued by Dercon et al. (2009) and Krishnan and Patnam (2014) unobserved heterogeneity must be considered when estimating the effect of training provision and participation. Formally, as described by Angrist and Pischke (2008), the observed difference in outcomes between the farmers who have been trained ($Y_{1i}|D_i = 1$) and those that have not ($Y_{0i}|D_i = 0$), is explained both by the causal effect of the training ($\kappa = Y_{1i} - Y_{0i}$) and a selection bias ($E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]$), the latter stemming from outcome-relevant, initial differences between the farming who where trained and those who were not:

$$\begin{aligned} & E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0] \\ & = \{\kappa + E[Y_{0i}|D_i = 1]\} - E[Y_{0i}|D_i = 0] \\ & = \kappa + \{E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]\} \end{aligned} \tag{2}$$

### iii. The Model

To identify the causal effects of learning from training and peers, we require several steps. In our case, the choice of the peer network is comparably straightforward, because the farmers are organized in local farmers' groups, which they also report as their main peer network (table 3 in the previous section).

This is only an approximation to the actual peer network, as most farmers are part of multiple, overlapping peer groups, such as neighbors, friends, and farmers' groups. However, the farmers' group is the main and most important network of the farmers, and discussing farming practices and business decisions is a main motivation for joining or starting a local farmers' group. In these groups, all members are pineapple farmers, and most external communication is organized through the farmers' groups.

Zeitlin (2011) develops an interesting approach to identify peer effects and uses data from Ghana's cocoa farmers, who are also organized in farmers' groups. However, the cocoa farmers' groups are local branches of larger organizations, which is a prerequisite for identification. In our case, most pineapple farmers' groups are truly local. As we are interested in a binary variable (adoption of mulching yes or no, adoption of organic fertilizers yes or no), we could use a discrete choice model, such as proposed by Brock and Durlauf (2001). Taking into account the endogeneity of peer learning and training is not trivial this way (Angrist 2001). As an example, Petrin and Train (2010) suggest the use of control functions, but they work better for training than for peer effects (because training is binary and peer learning is continuous). For continuous treatment variables, the Special Regressor approach of Lewbel et al. (2012) and Dong and Lewbel (2015) is feasible. However, as Angrist (2001) and Angrist and Pischke (2008) argue, the analytical framework can be much simplified by using a linear model and control for the endogeneity of explanatory variables with instrumental variables. This approach requires very little assumptions, is robust, and estimates are readily interpretable. It should be noted that 2SLS is simply the best linear approximation to the average treatment effect estimated with any discrete choice model that controls for endogeneity (Angrist and Pischke 2008). Thus, we choose a 2SLS framework for our analysis and we instrument training treatment and peer effects following the approach developed by Bramoullé et al. (2009), which has recently been employed by Krishnan and Patnam (2014) in a similar context to ours.

To further ease interpretability of our estimates, we always use standardized variables for the right hand side throughout this study, meaning that variables are rescaled to have a mean of zero and a standard deviation of one.

We begin our analysis then with an OLS regression, to establish that trainings are not systematically biased towards farmer more or less likely to adopt the trained innovation:

$$\begin{aligned} Tr_{izt+1} &= \beta_1 \overline{Tr}_{-izt} + \beta_2 x_{it} + \beta_3 F + \beta_4 y_{izt} + u_{izt} \quad if\ t\ \geq 2011 \\ Tr_{izt+1} &= \beta_1 \overline{Tr}_{-izt} + \beta_2 x_{it} + \beta_3 F + \beta_4 y_{izt} + u_{izt} \quad if\ t\ \leq 2011 \end{aligned} \tag{3}$$

Where $Tr_{izt}$ is whether or not a farmer $i$ is trained in a particular practice $z$ in period $t$+1 , $\overline{Tr}_{-izt}$ is the share of farmers that have been trained in her district so far (excluding herself), $x_{it}$ is a vector of explanatory variables that could affect the likelihood of being trained, $F$ is a vector of year and region dummies, and $y_{izt-1}$ depicts whether the farmer already used the innovation in the previous period. To understand whether the characteristics of the trained farmers change in time, we estimate equation 3 separately for earlier and later periods. It shows whether trainings were first offered to farmers more in need (e.g. less income, more constraints) or with a higher innovation potential (e.g. more income, less constraints) or whether possible trainings became more targeted in time, and also whether trained farmers are more or less likely to have already adopted the trained innovation before.

For this and all following models, we always estimate a few specifications, to probe the sensitivity of the estimates to the inclusion of various control variables. We usually start without any control variables, then proceed with the inclusion of strictly exogenous controls, and end with controls that are potentially endogenous. Standard errors are clustered at the farmers' group level, to take into account unobservables at this level.

For brevity and space, we only report our main results, less central results can be found in the online appendix to this article.

To turn to our main model, we need to consider the likely endogeneity of peer-learning and training, so we use the approach of Bramoullé et al. (2009), but adapt it to our available data. We first outline the approach and then describe how we apply it. The idea is that if peer behavior is endogenous, one can use the behavior of the peers of one's peers as instrument. This is possible if a network is characterized by a small degree of intransitivity (farmer $i$ is connected to farmer $j$, and farmer $j$ is connected to farmer $k$, but farmer $k$ is *not* connected to farmer $i$). The intuition for this instrument is that farmer $k$ can only affect farmer $i$ through affecting farmer $j$, so whatever farmer $i$ and farmer $j$ have in common (a common context, similar characteristics), farmer $i$ and farmer $k$ do not (they are not even connected).

The same instrument can be applied to instrument for training, but with a different rationale behind it. As we have described earlier, trainings move from place to place, so if there was a training organized in an adjacent community in the last period, chances are that farmers in close by communities will soon receive trainings themselves. We expect that trainers sometimes change their training schedule to offer a particular training in a particular place, and that farmers with a higher propensity to adopt an innovation are more likely to join a training. However, the share of training that is explained by the share of trained farmers in adjacent communities in the last period is free of community specific incentives and thus a feasible instrument.

As a caveat, we do not have detailed GPS data on the locations of the farms but only the names of the communities. The common approach would be to construct a neighbor matrix **W**, e.g. defined using the K nearest neighbors of a farmer, computed by the Euclidean distance between the farms, and to interact this matrix with the outcome variable and the exogenous peer characteristics. Instead, we rely on farmers' group and community locations, so that the lagged share of trained farmers in neighboring communities is our instrument for whether or not a farmer is trained and the lagged diffusion of an innovation in close by farmers' groups is our instrument for the diffusion of that innovation in a farmer's farmers' group. Our approach is thus the approach of Bramoullé et al. (2009) but done by hand. We control for correlated and selection effects by taking into account the behavior of the farmer in the last period, as well as past trainings and peer diffusion, as well as with dummies for location and period.

Our 2SLS specification then looks as follows:

$$y_{izt+1} = \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 \widehat{Tr}_{izt} + \beta_4 \hat{\bar{y}}_{-izt} + \beta_5 F + u_{it} \tag{4a}$$

$$\bar{y}_{-izt} = \alpha_1 \bar{y}_{jzt} + \alpha_2 \overline{Tr}_{jzt} + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \bar{y}_{-izt} + \beta_5 F + u_{it} \tag{4b}$$

$$Tr_{izt} = \alpha_1 \bar{y}_{jzt} + \alpha_2 \overline{Tr}_{jzt} + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \bar{y}_{-izt} + \beta_5 F + u_{it} \tag{4c}$$

Where $y_{it}$ denotes the technology choice of individual $i$ at time $t$ and $\bar{y}_{jzt}$ and $\overline{Tr}_{jzt}$ are our instrumental variables, namely the diffusion of an innovation amongst indirect neighbors (neighbors of neighbors) and the share of farmers training there. All other variables are defined as before. We continue to cluster the standard errors at the farmers' group level.

A variable that we have ignored so far is whether the farmer participates in contract farming. On the one hand, contract farming increases both the likelihood of being trained and also the ability and incentive to adopt innovations. On the other hand, contract farming is likely to be endogenous, so we do not want to naively enter contract farming as another control variable into the model. Instead, we estimate another 2SLS, in which we instrument whether or not the farmer has a farming contract with a company with the distance between the community of the farmer and the next company. The exclusion restriction is fulfilled because the companies are located at the center of the main pineapple production area, with sufficient distance to potential explanatory variables, such as the coast, the mountains, and especially the major cities (see figure 2 in section 2). Thus, our second 2SLS specification also controls for contract farming:

$$y_{izt+1} = \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 \widehat{Tr}_{izt} + \beta_4 \hat{\bar{y}}_{-izt} + \beta_5 \hat{C}_{it} + \beta_6 F + u_{it} \tag{5a}$$

$$\bar{y}_{-izt} = \alpha_1 \bar{y}_{jzt} + \alpha_2 \overline{Tr}_{jzt} + \alpha_3 dist_{it} + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \bar{y}_{-izt} + \beta_5 C_{it} + \beta_6 F + u_{it} \quad (5b)$$

$$Tr_{izt} = \alpha_1 \bar{y}_{jzt} + \alpha_2 \overline{Tr}_{jzt} + \alpha_3 dist_{it} + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \bar{y}_{-izt} + \beta_5 C_{it} + \beta_6 F + u_{it} \quad (5b)$$

$$C_{it} = \alpha_1 \bar{y}_{jzt} + \alpha_2 \overline{Tr}_{jzt} + \alpha_3 dist_{it} + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \bar{y}_{-izt} + \beta_5 C_{it} + \beta_6 F + u_{it} \quad (5b)$$

Finally, we estimate two more models, to learn a little more details of about estimates. First, we interact the model from equation 5 with period dummies, to estimate period specific effects of training and peer learning. Secondly, we split our sample into contract farmers and non-contract farmers, to see how much financial constraints matter for our results.

The period specific effects model looks as follows:

$$y_{izt+1} = \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} * Y + \beta_4 \overline{y}_{-izt} * Y + \beta_5 F + u_{it} \tag{6a}$$

$$\overline{y}_{-izt} * Y = \alpha_1 \overline{y}_{jzt} * Y + \alpha_2 \overline{Tr}_{jzt} * Y + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \overline{y}_{-izt} + \beta_5 F + u_{it} \tag{6b}$$

$$Tr_{izt} * Y = \alpha_1 \overline{y}_{jzt} * Y + \alpha_2 \overline{Tr}_{jzt} * Y + \beta_1 y_{izt} + \beta_2 x_{it} + \beta_3 Tr_{izt} + \beta_4 \overline{y}_{-izt} + \beta_5 F + u_{it} \tag{6c}$$

Where $Y$ are year dummies for the years 2009 to 2013.

The models on the subsamples looks exactly as equation 5a to 5c but is separately estimated for contract and non-contract farmers.

## 4. Results

We begin our analysis with a standard regression describing who are the farmers who receive and participate in training before and after 2011. The cut-off is arbitrarily set in the middle between 2009 and 2013, to be able to detect changes in the farmer characteristics. This is helpful later, when we consider time trends in the effectiveness of training. As table 3 shows, the best predictor for receiving training is how many other farmers have already been trained in the district. Since trainings move from location to location and close to all farmers participate, variables such as past adoption of the innovation, age, or education do not predict trainings. Variables that explain some share of the trainings are contract farming (because some trainings are offered by companies and companies can also help to organize training with other organizations), a higher share of nonfarm income (which could be a proxy for regional economic dynamism), and rainfall before 2011 but not later.

**Table 3. Who receives Training? (OLS)**

| spec | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| dv | training organic fertilizers | training organic fertilizers | training mulch | training mulch |
| period | <= 2011 | => 2011 | <= 2011 | => 2011 |
| district training | 1.288*** | 1.249*** | 1.432* | 1.087*** |
| | (0.253) | (0.335) | (0.777) | (0.399) |
| lag adoption | -0.190 | -0.204 | -0.0609 | 0.0184 |
| | (0.206) | (0.348) | (0.0952) | (0.0865) |
| age | 0.00910 | 0.00351 | 0.00372 | -0.0110 |
| | (0.0141) | (0.0173) | (0.0160) | (0.0182) |
| edu | -0.0201 | -0.0209 | 0.00594 | 0.0189 |
| | (0.0152) | (0.0168) | (0.0149) | (0.0188) |
| start | -0.0247 | -0.0258 | -0.0229 | -0.0501** |
| | (0.0175) | (0.0203) | (0.0221) | (0.0227) |
| contract | 0.0363 | 0.0443* | 0.0755*** | 0.103*** |
| | (0.0275) | (0.0255) | (0.0287) | (0.0237) |
| city | 0.00807 | 0.0170 | 0.0212 | 0.0391** |
| | (0.0150) | (0.0230) | (0.0179) | (0.0173) |
| company | -0.0338* | -0.0308 | -0.000195 | 0.0104 |
| | (0.0198) | (0.0226) | (0.0150) | (0.0243) |
| nonfarm | 0.0427* | 0.0519* | -0.00110 | -0.00563 |
| | (0.0237) | (0.0278) | (0.0207) | (0.0207) |
| rain | 0.0255* | 0.0114 | 0.0286** | 0.0141 |
| | (0.0148) | (0.0155) | (0.0117) | (0.0148) |
| rainvar | -0.0122 | -0.00153 | 0.0361 | 0.0135 |
| | (0.0216) | (0.0284) | (0.0221) | (0.0217) |
| soil | 0.0141 | 0.0148 | -0.00345 | -0.00805 |
| | (0.0142) | (0.0171) | (0.0151) | (0.0209) |
| R-sq | 0.18 | 0.17 | 0.11 | 0.16 |
| N | 936 | 936 | 936 | 936 |

Notes: The model is a random parameter OLS regression. Standard errors are clustered at the farmers' group level. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions with fixed effects (FE). To see whether the training target groups have changed in time, we split the sample in the year 2011. The full sample is 1990 observations.

In general, table 3 does not suggest that the characteristics of trained farmers has significantly changed in time, nor does it indicate a strong selection bias for training. Nevertheless, we control for a selection bias in both peer learning and training, because of the risk that selection occurs based on omitted variables.

**Table 4. Adoption of Organic Practices Second Stage (2SLS)**

| spec | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| adoption of | org fert | org fert | org fert | mulch | mulch | mulch |
| training | 0.0361** | 0.0348** | 0.0351** | 0.0196 | 0.0203 | 0.0177 |
| | (0.0150) | (0.0146) | (0.0146) | (0.0131) | (0.0132) | (0.0136) |
| peer | 0.0374*** | 0.0373*** | 0.0373*** | 0.190*** | 0.188*** | 0.188*** |
| | (0.0134) | (0.0135) | (0.0134) | (0.0203) | (0.0212) | (0.0209) |
| rain | | 0.0101* | 0.0101* | | -0.0129 | -0.0131 |
| | | (0.00605) | (0.00605) | | (0.0117) | (0.0117) |
| farmsize | | 0.00743** | 0.00789** | | 0.0196** | 0.0158* |
| | | (0.00513) | (0.00520) | | (0.0133) | (0.0131) |
| risk pref | | 0.00430 | 0.00431 | | -0.0126 | -0.0129 |
| | | (0.00412) | (0.00413) | | (0.0128) | (0.0126) |
| nonfarm | | -0.00318 | -0.00310 | | 0.00418 | 0.00315 |
| | | (0.00315) | (0.00306) | | (0.00827) | (0.00838) |
| credit | | | -0.00188 | | | 0.0151 |
| | | | (0.00394) | | | (0.0113) |
| controls | A | B | C | A | B | C |
| R-sq | 0.79 | 0.79 | 0.79 | 0.50 | 0.50 | 0.51 |
| F train | 65.77 | 94.22 | 100.17 | 479.64 | 493.72 | 387.92 |
| F peer | 2385.34 | 2192.19 | 2423.42 | 298.49 | 302.22 | 307.50 |
| model | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets (clustered at the farmers' group). F train is the Craig Donald F value for the excluded instrument for training (the training of indirect neighbors) and F peers is the Craig Donald F value for the excluded instrument for peer-learning (the innovation diffusion amongst indirect neighbors). Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

Table 4 presents our estimates of the effectiveness of training and peer learning. The first stages are presented in the appendix. Below the table we show the Craig Donald F-values which suggests that the share of trained neighbors and the technology diffusion amongst them are strong instruments for a farmers training participation and opportunity for peer learning. It is suggested that training and peer learning are very similarly effective to diffuse organic fertilizers (org fert) – at a rather low level (about 4% increase in the probability that organic fertilizers are adopted). In contrast, training does not significantly increase the probability that mulching is adopted, but the main driver is peer-learning, which is very effective (with a 19% increase in the probability that mulching is adopted). A few other variables are shown for comparison. It can be seen that rainfall is important for the adoption of organic fertilizers (because many organic fertilizers are living plants that require water) but not for mulching (because mulching materials are non-living). Farmsize is significant for both technologies, but the effect is small. Neither nonfarm income, risk preferences, or credit is estimated to be important for the adoption of the analyzed technologies.

**Table 5. Adoption of Organic Practices Second Stage (2SLS)**

| spec<br>adoption of | (1)<br>org fert | (2)<br>org fert | (3)<br>org fert | (4)<br>mulch | (5)<br>mulch | (6)<br>mulch |
|---|---|---|---|---|---|---|
| training | 0.0327** | 0.0314** | 0.0306** | -0.0205 | -0.0327 | -0.0346 |
| | (0.0119) | (0.0119) | (0.0119) | (0.0216) | (0.0236) | (0.0234) |
| peer | 0.0335** | 0.0327* | 0.0313* | 0.166*** | 0.157*** | 0.157*** |
| | (0.0110) | (0.0121) | (0.0125) | (0.0160) | (0.0172) | (0.0174) |
| contract | 0.0213 | 0.0217 | 0.0278 | 0.111 | 0.152* | 0.152* |
| | (0.0172) | (0.0211) | (0.0237) | (0.0510) | (0.0597) | (0.0605) |
| controls | A | B | C | A | B | C |
| R-sq | 0.79 | 0.79 | 0.79 | 0.48 | 0.46 | 0.46 |
| F train | 316.10 | 330.39 | 333.22 | 623.99 | 848.19 | 717.73 |
| F peer | 444.23 | 341.64 | 336.21 | 863.70 | 1098.69 | 1121.65 |
| F contract | 23.16 | 24.12 | 22.79 | 45.53 | 52.24 | 51.27 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets (clustered at the farmers' group). F train is the Craig Donald F value for the excluded instrument for training (the training of indirect neighbors), F peers shows the same for the excluded instrument for peer-learning (the innovation diffusion amongst indirect neighbors), and F contract shows this for contract farming (the instrument is the distance to the closest company). Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

Whereas the results of table 4 are suggestive, we need to consider the effect of contract farming. Contract

farming could both increase the chance of receiving training and incentivize the adoption of new technologies.

However, we cannot naively control for contract farming because of a likely selection bias in who becomes a contract farmer and who does not. To exogenize contract farming, we use the distance to the closest pineapple processing company as an instrument. The exclusion restriction is fulfilled because the distance to the companies does not much correlate with any other relevant location, such as distance to the coast or the capital. Below table 5, we show the Craig Donald F values again – the full first stage results can be seen in the appendix. The results of table 5 suggest that contract farming is only a significant adoption determinant for mulching but not for organic fertilizers. For mulching, however, the estimated effect is large and comparable to the effect of peer-learning. The basic pattern from table 4, however, does not change: Training is a significant adoption determinant for organic fertilizers and mulching is mostly learned from the peers.

To better understand the results above, we consider two more tests. First, we estimate period specific effects for training and peer learning, to see whether there are obvious trends in our data. Secondly. We split our sample into farmers who currently have a farming contract with a company and those who do not, to investigate whether the effects of training and peer learning are distinct for these two groups (mulching, e.g. is more expensive than organic fertilizers, so we might expect contract farmers to be less constrained to adopt mulching than other farmers).

Beginning with table 6, we do not see clear time trends between 2009 and 2013. Our basic observed pattern is relatively stable over time, which is likely the result of our relatively short time frame.

More informative is table 7. We can see that training is generally more effective for contract farmers, which could be due to the individual characteristics of these farmers or because of the complementary benefits that contract farming is providing. For the use of organic fertilizers, training increases the adoption probability by about 7% for contract farmers and 3% for non-contract farmers. For mulching, the adoption probability is only significantly increased for contract farmers, but the estimated effect is still very much smaller than the effect of peer learning

(3% versus 19%). These results suggest that farmer characteristics play a role in the effectiveness of the provided trainings, but this role is only minor in comparison to the effect of innovation characteristics.

**Table 6. Period Specific Effects (2SLS)**

| spec<br>adoption of | (1)<br>org fert | (2)<br>org fert | (3)<br>org fert | (4)<br>mulch | (5)<br>mulch | (6)<br>mulch |
|---|---|---|---|---|---|---|
| training 09 | 0.00807 | 0.00661 | 0.00670 | -0.00587 | -0.00588 | -0.00546 |
| | (0.0262) | (0.0258) | (0.0259) | (0.0316) | (0.0324) | (0.0322) |
| training 10 | 0.0493*** | 0.0480*** | 0.0483*** | 0.0275 | 0.0293 | 0.0289 |
| | (0.0157) | (0.0155) | (0.0155) | (0.0280) | (0.0277) | (0.0277) |
| training 11 | 0.0384*** | 0.0363*** | 0.0366*** | 0.0290 | 0.0291 | 0.0293 |
| | (0.0124) | (0.0119) | (0.0119) | (0.0306) | (0.0315) | (0.0310) |
| training 12 | 0.0319* | 0.0301* | 0.0306* | 0.0353 | 0.0360 | 0.0313 |
| | (0.0185) | (0.0182) | (0.0183) | (0.0218) | (0.0222) | (0.0229) |
| training 13 | 0.0475*** | 0.0473*** | 0.0476*** | 0.00532 | 0.00546 | -0.000384 |
| | (0.0173) | (0.0172) | (0.0172) | (0.0186) | (0.0189) | (0.0198) |
| peer 09 | 0.0279* | 0.0247 | 0.0249 | 0.201*** | 0.196*** | 0.195*** |
| | (0.0158) | (0.0153) | (0.0153) | (0.0280) | (0.0277) | (0.0281) |
| peer 10 | 0.0323** | 0.0308** | 0.0306** | 0.192*** | 0.187*** | 0.189*** |
| | (0.0135) | (0.0141) | (0.0141) | (0.0200) | (0.0197) | (0.0199) |
| peer 11 | 0.0289 | 0.0291 | 0.0289 | 0.183*** | 0.181*** | 0.182*** |
| | (0.0198) | (0.0197) | (0.0197) | (0.0198) | (0.0203) | (0.0204) |
| peer 12 | 0.0258 | 0.0258 | 0.0257 | 0.182*** | 0.180*** | 0.180*** |
| | (0.0184) | (0.0182) | (0.0181) | (0.0180) | (0.0185) | (0.0183) |
| peer 13 | 0.0526** | 0.0532** | 0.0535** | 0.191*** | 0.191*** | 0.193*** |
| | (0.0246) | (0.0245) | (0.0246) | (0.0173) | (0.0178) | (0.0177) |
| controls | A | B | C | A | B | C |
| R-sq | 0.79 | 0.79 | 0.79 | 0.51 | 0.51 | 0.51 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. The latter are clustered at the community and year level. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 7. How much do Financial Incentives and Constraints Matter? (2SLS)**

| spec | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| adoption of | organic fertilizers | organic fertilizers | mulch | mulch |
| contract | no | yes | no | yes |
| training | 0.0264*** | 0.0743*** | 0.00732 | 0.0289* |
| | (0.00835) | (0.0135) | (0.0179) | (0.0149) |
| group | 0.0381*** | 0.0427*** | 0.185*** | 0.192*** |
| | (0.00528) | (0.00987) | (0.0146) | (0.0225) |
| controls | B | B | B | B |
| F excl. 1 | 766.82 | 132.78 | 638.87 | 340.10 |
| F excl. 2 | 1676.24 | 838.35 | 3076.48 | 546.58 |
| R-sq | .75 | .83 | .47 | .41 |
| N | 1425 | 565 | 1425 | 565 |

Notes: The table reports estimated coefficients and standard errors in brackets. The latter are clustered at the group level. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE).

## 5. Discussion and Conclusion

The increased use of sustainable intensification practices (such as mulching and organic fertilizers) would be highly beneficial for the pineapple farmers in Ghana (World Bank 2011, Kleemann and Abdulai 2013). Because such technologies are knowledge intensive, they are mostly promoted through trainings, provided by development organizations, extension agents (private and governmental), and NGOs. Above, we find that the degree of complexity is the main explanation why mulching is much wider diffused than organic fertilizers. The very specific knowledge and adaptations required by organic fertilizers makes peer learning comparably ineffective. This is consistent with the model of Munshi (2004), who finds that peer learning can be constrained by unobserved population heterogeneity. Munshi (2004) considers the diffusion of new rice and wheat varieties during the Indian revolution and heterogeneity comes from differences between the regions in which rice and wheat are grown (larger heterogeneity in the rice growing region). We show that how much such heterogeneity matters is technology dependent. In contrast to the view that all sustainable intensification technologies are complex and knowledge intensive, we make the point that there is actually a spectrum and mulching is at the simpler end. This makes peer-learning effective and once the technology has started to diffuse, no more training is needed. This is consistent with the finding of Krishnan and Patnam (2014), who investigate the comparative effects of training and peer-learning for the diffusion of new seeds and chemical fertilizer in Ethiopia and who find that training is more effective to start the diffusion process than to enhance it later. However, we find that this depends on the nature of the innovation, as relatively simple innovations such as conventional inputs or mulching easily diffuse through peer learning but more complex innovations such as integrating new crops (into the rotation or to plant them together with the main crop) require more specific knowledge and thus profit from training significantly more.

Our results suggest that in order to diffuse sustainable intensification practices more widely, the focus should be put on organic fertilizers, as mulching is going to diffuse through peer learning on its own. To achieve a similar diffusion of organic fertilizers (about 50% adopters), training on the use of organic fertilizers must likely more than double (from 10% to 25%).

# Appendix

In the following we present a few tables that did not make it into the main text. We start with table 4a and 4b, which are the first stage estimates for table 4 in the main text. Table 4c shows the estimates for the second stage but in contrast to the table in the main text, we also show standard errors that are clustered at the farmer level, for comparison. Tables 5a, 5b, and 5c have the same purpose but for table 5 in the main text. Table 6a shows the first stage Craig Donald F-values for table 6 in the main text (period specific effects).

**Table 4a. Adoption of Organic Practices First Stage organic fertilizers (2SLS)**

| spec | (1) | (1) | (2) | (2) | (3) | (3) |
|---|---|---|---|---|---|---|
| dv | training | group | training | group | training | group |
| n_train | 0.651*** | 0.0987** | 0.656*** | 0.101** | 0.655*** | 0.101*** |
| | (0.0618) | (0.0379) | (0.0599) | (0.0387) | (0.0589) | (0.0384) |
| n_adopt | -0.0305* | 0.767*** | -0.0441** | 0.766*** | -0.0461** | 0.768*** |
| | (0.0164) | (0.0120) | (0.0196) | (0.0125) | (0.0191) | (0.0121) |
| controls | A | A | B | B | C | C |
| R-sq | 0.73 | 0.77 | 0.73 | 0.77 | 0.74 | 0.77 |
| Craig Donald F | 65.77 | 2385.34 | 94.22 | 2192.19 | 100.17 | 2423.42 |
| model | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. For brevity, only the group level clustered standard errors are reported. Farmer level clustered standard errors can be obtained from the authors upon request. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 4b. Adoption of Organic Practices First Stage Mulching (2SLS)**

| spec | (1) | (1) | (2) | (2) | (3) | (3) |
|---|---|---|---|---|---|---|
| dv | training | group | training | group | training | group |
| n_train | 0.775*** | 0.0224 | 0.779*** | 0.0230 | 0.762*** | 0.0264 |
| | (0.0253) | (0.0180) | (0.0249) | (0.0169) | (0.0274) | (0.0171) |
| n_adopt | -0.0468** | 0.787*** | -0.0520*** | 0.782*** | -0.0488** | 0.781*** |
| | (0.0180) | (0.0331) | (0.0193) | (0.0324) | (0.0201) | (0.0320) |
| controls | A | A | B | B | C | C |
| R-sq | 0.60 | 0.85 | 0.61 | 0.85 | 0.62 | 0.86 |
| Craig Donald F | 479.64 | 298.49 | 493.72 | 302.22 | 387.92 | |
| model | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. For brevity, only the group level clustered standard errors are reported. Farmer level clustered standard errors can be obtained from the authors upon request. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 4c. Adoption of Organic Practices Second Stage (2SLS)**

| spec | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| adoption of | Organic fertilizers | Organic fertilizers | Organic fertilizers | Mulch | Mulch | Mulch |
| training | 0.0361** | 0.0348** | 0.0351** | 0.0196 | 0.0203 | 0.0177 |
| | (0.00670) | (0.00666) | (0.00667) | (0.0108) | (0.0108) | (0.0112) |
| | (0.0150) | (0.0146) | (0.0146) | (0.0131) | (0.0132) | (0.0136) |
| group | 0.0374*** | 0.0373*** | 0.0373*** | 0.190*** | 0.188*** | 0.188*** |
| | (0.00481) | (0.00493) | (0.00492) | (0.0119) | (0.0122) | (0.0122) |
| | (0.0134) | (0.0135) | (0.0134) | (0.0203) | (0.0212) | (0.0209) |
| rain | | 0.0101* | 0.0101* | | -0.0129 | -0.0131 |
| | | (0.00408) | (0.00408) | | (0.00956) | (0.00955) |
| | | (0.00605) | (0.00605) | | (0.0117) | (0.0117) |
| farmsize | | 0.00743** | 0.00789** | | 0.0196** | 0.0158* |
| | | (0.00349) | (0.00361) | | (0.00812) | (0.00843) |
| | | (0.00513) | (0.00520) | | (0.0133) | (0.0131) |
| risk pref | | 0.00430 | 0.00431 | | -0.0126 | -0.0129 |
| | | (0.00388) | (0.00388) | | (0.00925) | (0.00925) |
| | | (0.00412) | (0.00413) | | (0.0128) | (0.0126) |
| nonfarm | | -0.00318 | -0.00310 | | 0.00418 | 0.00315 |
| | | (0.00358) | (0.00359) | | (0.00814) | (0.00816) |
| | | (0.00315) | (0.00306) | | (0.00827) | (0.00838) |
| credit | | | -0.00188 | | | 0.0151 |
| | | | (0.00368) | | | (0.00894) |
| | | | (0.00394) | | | (0.0113) |
| controls | A | B | C | A | B | C |
| R-sq | 0.79 | 0.79 | 0.79 | 0.50 | 0.50 | 0.51 |
| model | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS | 2SLS |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. The upper brackets show the standard errors of a random parameter OLS regression, the lower brackets show standard errors that are clustered at the group level. When the significance differed between the two models, the stars are assigned according to the lower significance. Levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 5a. Adoption of Organic Practices First Stage Organic Fertilizers (2SLS)**

| spec | (1) | (1) | (1) | (2) | (2) | (2) | (3) | (3) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| dv | training | group | contract | training | group | contract | training | group | contract |
| n_train | 0.649*** | 0.0996*** | 0.130*** | 0.655*** | 0.101*** | 0.134*** | 0.654*** | 0.102*** | 0.131*** |
| | (0.0220) | (0.0213) | (0.0360) | (0.0213) | (0.0213) | (0.0342) | (0.0213) | (0.0214) | (0.0332) |
| n_adopt | -0.0288* | 0.766*** | 0.127*** | -0.0420** | 0.765*** | 0.142*** | -0.0438*** | 0.767*** | 0.137*** |
| | (0.0154) | (0.0222) | (0.0326) | (0.0165) | (0.0248) | (0.0339) | (0.0166) | (0.0250) | (0.0339) |
| distance | 0.0329** | -0.0220*** | -0.207*** | 0.0283** | -0.0156** | -0.189*** | 0.0315** | -0.0179** | -0.180*** |
| | (0.0128) | (0.00578) | (0.0381) | (0.0135) | (0.00700) | (0.0368) | (0.0136) | (0.00735) | (0.0353) |
| R-sq | 0.73 | 0.77 | 0.12 | 0.73 | 0.77 | 0.19 | 0.74 | 0.77 | 0.20 |
| F excl. | 316.10 | 444.23 | 23.16 | 330.39 | 341.64 | 24.12 | 333.22 | 336.21 | 22.79 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. For brevity, only the group level clustered standard errors are reported. Farmer level clustered standard errors can be obtained from the authors upon request. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

## Table 5b. Adoption of Organic Practices First Stage Mulching (2SLS)

| spec | (1) | (1) | (1) | (2) | (2) | (2) | (3) | (3) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| dv | training | group | contract | training | group | contract | training | group | contract |
| n_train | 0.775*** | 0.0214* | 0.275*** | 0.778*** | 0.0219* | 0.266*** | 0.765*** | 0.0249** | 0.256*** |
| | (0.0165) | (0.0116) | (0.0318) | (0.0165) | (0.0116) | (0.0303) | (0.0176) | (0.0118) | (0.0293) |
| n_adopt | -0.0464*** | 0.788*** | 0.158*** | -0.0516*** | 0.783*** | 0.140*** | -0.0490*** | 0.782*** | 0.142*** |
| | (0.0136) | (0.0138) | (0.0322) | (0.0141) | (0.0137) | (0.0313) | (0.0138) | (0.0135) | (0.0313) |
| distance | -0.00757 | -0.0181 | -0.200*** | -0.0152 | -0.0217 | -0.177*** | -0.00827 | -0.0233* | -0.172*** |
| | (0.0148) | (0.0132) | (0.0302) | (0.0150) | (0.0135) | (0.0308) | (0.0143) | (0.0136) | (0.0300) |
| R-sq | 0.60 | 0.85 | 0.22 | 0.61 | 0.86 | 0.26 | 0.62 | 0.86 | 0.27 |
| F excl. | 623.99 | 863.70 | 45.53 | 848.19 | 1098.69 | 52.24 | 717.73 | 1121.65 | 51.27 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. For brevity, only the group level clustered standard errors are reported. Farmer level clustered standard errors can be obtained from the authors upon request. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 5c. Adoption of Organic Practices Second Stage (2SLS)**

| spec | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| adoption of | ORGANIC FERTILIZERS | ORGANIC FERTILIZERS | ORGANIC FERTILIZERS | Mulch | Mulch | Mulch |
| training | 0.0327** | 0.0314** | 0.0306** | -0.0205 | -0.0327 | -0.0346 |
| | (0.00733) | (0.00758) | (0.00765) | (0.0223) | (0.0255) | (0.0254) |
| | (0.0119) | (0.0119) | (0.0119) | (0.0216) | (0.0236) | (0.0234) |
| group | 0.0335** | 0.0327* | 0.0313* | 0.166*** | 0.157*** | 0.157*** |
| | (0.00630) | (0.00721) | (0.00730) | (0.0167) | (0.0183) | (0.0186) |
| | (0.0110) | (0.0121) | (0.0125) | (0.0160) | (0.0172) | (0.0174) |
| contract | 0.0213 | 0.0217 | 0.0278 | 0.111 | 0.152* | 0.152* |
| | (0.0207) | (0.0240) | (0.0253) | (0.0536) | (0.0653) | (0.0664) |
| | (0.0172) | (0.0211) | (0.0237) | (0.0510) | (0.0597) | (0.0605) |
| | (0.0168) | (0.0168) | (0.0168) | (0.0222) | (0.0231) | (0.0232) |
| | (0.0380) | (0.0379) | (0.0377) | (0.0277) | (0.0289) | (0.0291) |
| controls | A | B | C | A | B | C |
| R-sq | 0.79 | 0.79 | 0.79 | 0.48 | 0.46 | 0.46 |
| N | 1990 | 1990 | 1990 | 1990 | 1990 | 1990 |

Notes: The table reports estimated coefficients and standard errors in brackets. The upper brackets show the standard errors of a random parameter OLS regression, the lower brackets show standard errors that are clustered at the group level. When the significance differed between the two models, the stars are assigned according to the lower significance. Levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

**Table 6a. F Values of the Excluded Instruments**

| spec | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| training 1 | 221.70 | 228.43 | 234.10 | 15.22 | 20.09 | 19.96 |
| training 2 | 297.04 | 292.08 | 293.26 | 29.94 | 33.45 | 33.42 |
| training 3 | 232.49 | 256.36 | 259.27 | 38.24 | 43.15 | 43.13 |
| training 4 | 254.59 | 269.60 | 264.02 | 63.38 | 68.88 | 65.96 |
| training 5 | 106.25 | 109.04 | 109.08 | 64.42 | 68.76 | 65.37 |
| peer 1 | 9.71 | 12.65 | 12.81 | 48.11 | 51.07 | 50.01 |
| peer 2 | 14.31 | 19.23 | 19.30 | 59.05 | 62.34 | 62.35 |
| peer 3 | 15.18 | 17.58 | 17.66 | 53.39 | 53.54 | 53.64 |
| peer 4 | 12.86 | 18.34 | 17.70 | 54.44 | 55.01 | 55.62 |
| peer 5 | 44.87 | 68.13 | 67.28 | 65.05 | 65.17 | 65.39 |

Notes: The table reports estimated coefficients and standard errors in brackets. The latter are clustered at the community and year level. Significance levels are 10% (*), 5% (**), and 1% (***). We control for unobservable differences between the regions and years with fixed effects (FE). The specifications differ by their set of control variables. Set A includes only the lagged adoption of each farmer. Set B also includes rainfall, soil quality, age, education, farmsize, risk preference, and nonfarm income. Set C additionally includes whether the farmer received a credit.

# Chapter 6

# The World Heritage List: Which Sites Promote the Brand? A Big Data Spatial Econometrics Approach

with Marc Patry[7]

**Abstract** UNESCO's World Heritage Convention encourages inscribed sites to promote the World Heritage brand by clearly communicating their affiliation. Based on feedback from over 319,000 visitors at 791 locations, we create an index that shows the extent to which World Heritage sites are actually branding themselves as such. We find great heterogeneity throughout the list and explain this econometrically with site specific incentives. Notably, the sites that benefit more from the World Heritage brand are significantly more willing to contribute to the collective brand than sites that benefit less. Specifically, rural sites are much better branded than urban sites, sites, as rural sites benefit more from the brand than urban sites. We also find a positive relationship between World Heritage branding and its conservation status and a U-shaped relationship between a site's visitor numbers and its branding. Furthermore, Asian sites are much better branded than sites in the Middle East, and richer countries and those with already more international tourists are branded less. The difficulty of effective branding, e.g. for large, open access sites, has no significant effect. Our findings suggest that mandatory World Heritage branding obligations would have a positive effect on the World Heritage brand equity, bringing conservation and economic benefits to a much wider range of World Heritage sites.

This chapter is published as Wuepper, D. and M. Patry (2016). "The World Heritage list: Which sites promote the brand? A big data spatial econometrics approach." In the Journal of Cultural Economics as online first.

Contributions Marc Patry provided the data as well as ideas and feedback to the article. David Wuepper came up with the research question, conducted the analysis and wrote the article.

[7] United Nations Educational, Scientific and Cultural Organization

# I. Introduction

With the World Heritage (WH) Convention, the United Nations Educational, Scientific and Cultural Organization (UNESCO) created an instrument that engages the international community in an effort to identify and conserve the worlds' most outstanding natural and cultural treasures and give these sites a function in the life of the communities in which they are embedded (UNESCO 1972).

The Convention also supports the development of a collective brand for the so far 1007 WH sites (as of January 2015) in 161 countries, which means it is a common interest of UNESCO and the WH sites to create collective brand equity (Mas and Nicolau 2010, Nicolau and Mas 2014) - i.e. to attract culture tourists, convince donors, strengthen political support, raise awareness about the importance of conservation and/or restauration.

Formally, brand equity is "a set of assets such as name awareness, loyal customers, perceived quality, and associations that are linked to the brand (its name and symbol) and add (or subtract) value to the product or service being added" (Aaker 2009). It is created mainly by the interplay of the value the WH sites are delivering and how much they are recognized as being a WH site.

In this view, it is the most iconic and popular sites that have the greatest ability to increase WH brand equity. From a political economy view however, it is questionable whether these sites actually use their ability to the fullest.

Marcotte and Bourdeau (2012) and King and Halpenny (2014) analyze the WH branding of WH sites in Israel, Australia and the USA and find that none of them are properly branded as WH sites and most visitors are completely unaware of the sites' status. Marcotte and Bourdeau (2012) investigate this matter online and find that most Western European cities are branded as WH sites online (mostly to attract tourists).

A likely explanation why some sites do not contribute to the collective equity of the WH brand is heterogeneity in the individual benefit of doing so, e.g. because some sites can increase their tourism income with the WH brand while others cannot (Jones and Munday 2001, Smith 2002, Tisdell and Wilson 2002, Buckley 2004, Yang et al. 2010, Jimura 2011, King 2011, Huang et al. 2012, Kayahan and Vanblarcom 2012, Hardiman and Burgin 2013).

Rural destinations, e.g., are more likely to benefit from WH branding than urban sites. The reason is that it is not risky for tourists to visit the historic city centers in Paris or Rome: First, the visitors know what they will experience and secondly, reaching the destination is simple and inexpensive. Rural destinations in contrast, are more risky: The visitors must often travel far to reach the site and there are few other attractions close by if the site disappoints. Thus, the inscription on the WH list makes potential visitors aware of less famous destinations and signals quality, thereby making a visit more attractive. Other factors that plausibly affect the WH tourism benefit are the marketability of different kind of sites as WH and the cultural background of visitors and managers. With the specific kind of site we mean that it is perhaps easier to promote a historic Viking settlement as WH than it is to promote a part of a larger city that is probably associated with many, competing associations. With cultural background we mean that in Asia, e.g., destinations are especially proud to be listed as WH and visitors want to experience such destinations. In the middle east, in contrast, historic Medinas and other heritage sites are considered national, Arabic, or Muslim heritage but not WH. Thus, managers are not particularly proud about WH inscription and visitors often do not even know about the WH list.

As Ostrom (1990) has pointed out, heterogeneous interests can make collective action - such as collective brand building – difficult if the stakeholders only take into account their own disparate interests. Furthermore, cultural goods have peculiarities that may favor the influence of special interest groups, such as the difficulty of evaluating the costs and benefit of specific public investments into culture or the public control over culture policies (Mazza 2011). Thus, it might be that local, national and international interests are not fully aligned and special interest groups have the scope to capture rents, i.e. by investing heavily in the promotion of places with high political or economic returns and to elude investments for places with low political or economic returns, largely ignoring the global public interest to promote and conserve humanity's heritage irrespective of local and short-term gains.

Indeed, analyzing the marketing efforts of WH destinations in Spain, Mas and Nicolau (2010) and Nicolau and Mas (2014) find that local interests dominate collective interests and sites try to free-ride on the marketing efforts of others. This suggest that not all destination currently profit from working towards the collective benefits of a stronger WH brand. Better understanding what drives on-site marketing behavior thus might contribute to make policy adjustment to align private and collective interests (e.g. by communicating best practices amongst similar sites, to enable more destinations to profit from their WH listing).

Related research by Bertacchini and Saccone (2012) and Frey et al. (2011) find that even earlier in the WH listing process, during the site selection onto the WH list, economic and political incentives motivate decisions, when in theory, it should only be a location's global public good character (Kaul et al. 1999) that should matter.

In summary, while the purpose of UNESCOs WH list is to conserve and promote various global public goods, empirical research suggests that economic and political stakeholders pursue their own interests and often use the WH brand for their individual rather for collective goals. In this context, however, the great heterogeneity of the WH list must be acknowledged: There are different motives why locations are submitted to become WH sites (Rebanks 2009) and this might affect how much the management of the inscribed sites focuses on global benefits versus site specific benefits.

In the following, we are interested in which sites promote the WH brand on-site by informing visitors about the association and its meaning. Is it the sites with the highest potential to do so (i.e. the most popular sites)? Or those with the highest ability (i.e. the most compact sites with single entry points)? Or those with the highest individual incentive (i.e. more remote, further away from the next city – to convince visitors that it is worth coming)?

Our theoretical framework is derived from microeconomic theory but with the non-standard adjustment that WH sites are companies with several goals and not just profit maximization. These goals include conservation, financial sustainability, and education, and they relate to market, non-market and cultural values (Throsby et al. 2012). It has been established that cultural value cannot fully be captured by economic value, and is thus separate even though there is a correlation (Hutter and Throsby 2008, Throsby and Zednik 2013). The difference is that economic value is commonly defined as the willingness to pay (WTP) of individuals, be it because of use (e.g. recreation) or non-use values (e.g. a preference for conservation), whereas cultural value is concerned with additional dimensions (societal values, or non-quantifiable values such as spiritual, historical or symbolic values). The WH Convention has the official goal to take into consideration the total value of the WH sites, economic market and non-market values, as well and especially, the cultural values. Whether this is reflected in observed management choices is investigated below.

The contribution of our study is hence twofold:

First, we collect online-survey data from a representative sample of 319,000 visitors at 791 WH sites to understand how well the individual locations are branded as WH sites *on-site*. Second, we then econometrically model the determinants of branding a location as a WH site to find out, whether UNESCO's encouragement to promote the brand on-site (UNESCO 2013) is sufficient to achieve high levels of WH branding or not.

We make the following findings: The sites use the WH brand for many different goals. Some sites use it to attract more visitors, other use it to increase support for conservation. How much a sites benefits from the WH brand explains well how much it promotes the WH brand on-site, and this benefit is often tourism income. Variables that determine the benefit from being a WH site include world region and countries, as preferences for the brand vary geographically. They also include the gdp of the country and its tourism popularity. Site specifically, rural sites benefit more than urban ones, sites with especially many or few visitors benefit more than sites with intermediate visitor numbers, and there is a positive correlation between the WH benefit and the conservation status of a site.

In the following we describe our data and method (2) our model (3) and our descriptive findings (4). Then, we present our analytical findings (5), followed by robustness Checks (6), a discussion (7) and a conclusion with policy recommendations (8).

## II. Data Collection and Method

In a partnership agreement between UNESCO's World Heritage Center (WHC) and TripAdvisor (TA), we collected feedback from more than 319.000 World Heritage visitors about their awareness of the WH logo, brochures or other information that announced the site's WH status. This feedback is the basis of our measure how much the WH sites are branded as such.

The data was collected online between 11/2009 and 11/2011 through the TA website. The purpose of this website is to provide people with a platform where they can give feedback about tourism related locations. In our case, whenever a person rated a hotel, restaurant or attraction near or within a WH site, they were prompted to answer the following additional questions:

"Did you know that XX is a World Heritage site?", "How good is the current condition of the site?" and "How aware were you of signs, plaques, pamphlets and other materials identifying the location as a World Heritage site?". The ratings were presented on a 1 to 5 Likert scale, where 1 indicated "not aware at all" for the visibility question and "very bad condition" for the condition question. On the other end, 5 indicated "very aware" for the visibility question and "very good condition" for the condition question. The respondents were also given the opportunity to comment their answers.

**Table 1: Sampling Representativeness**

| | full list | | | | sample 1 | | | | sample 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Regions | Cult | Nat | Mix | % | Cult | Nat | Mix | % | Cult | Nat | Mix | % |
| Africa | 48 | 37 | 4 | **9** | 15 | 12 | 1 | **5** | 10 | 6 | 1 | **4** |
| Arab States | 71 | 4 | 2 | **8** | 32 | 0 | 0 | **6** | 25 | 0 | 0 | **6** |
| Asia-Pacific | 161 | 59 | 11 | **23** | 89 | 27 | 0 | **21** | 62 | 14 | 0 | **18** |
| Europe and NA | 408 | 61 | 10 | **48** | 273 | 28 | 11 | **55** | 214 | 19 | 8 | **58** |
| Latin America | 91 | 36 | 4 | **13** | 62 | 11 | 2 | **13** | 44 | 8 | 1 | **13** |
| % | **77** | **20** | **3** | | **84** | **14** | **2** | | **86** | **11** | **2** | |

Because all our respondents are users of the TA-website, the representativeness of our sample depends on the representativeness of the TA-users. Because of TA's popularity, this is arguably fulfilled. Most importantly, we aggregate the feedback for each site in order to get our variables. Thus, only if TA-users would exhibit a significantly distinct rating pattern of the WH visibility compared to non-TA-users would our estimates loose generalizability.

The visitor feedback was then sent to the WHC on a monthly basis and stored in a database together with basic information about each site (provided by UNESCO's WH Centre), such as inscription date, size, whether or not it has ever been on UNESCO's danger list, the site's category, its world region and country, as well as each site's geographic coordinates.

To better understand what determines how much a location is branded as WH, we also collected additional data about the sites – especially on the benefit a location can obtain from WH branding. Together with the United Nations Volunteering Service, we collected the annual visitation numbers of the WH sites. These numbers are clearly imprecise and subject to significant measurement error. However, to the best of our knowledge, this is the first time, this information is collected at all and thus, having an imprecise number is certainly better than having no number at all. We carefully checked the data for obvious biases. We also used Arcgis software to map the WH sites on a global map and create a range of spatial variables such as distances to the coast, to mountains, to the closest mountain or coast, to other WH sites and to cities of different sizes, as well as variables that contain the number of features within different radii, such as the number of

### Table 2: Variables in the Econometric Framework

| variable | description | source |
|---|---|---|
| visibility | Index from 0 to 5 how well the site is branded | Own survey |
| C1 masterpiece | Site represents a masterpiece of human creative genius | UNESCO |
| C2 interchange | Site exhibits an important interchange of human values | UNESCO |
| C3 past tradition | Site bears a unique testimony to a tradition | UNESCO |
| C4 history | Site illustrates a significant stage of human history | UNESCO |
| C5 interaction | Outstanding example of human-environmental interaction | UNESCO |
| C6 living tradition | Site is associated with living traditions | UNESCO |
| N7 phenomenon | Site contains superlative natural phenomena | UNESCO |
| N8 history | Site represents a major stage of earth's history | UNESCO |
| N9 processes | Site is an outstanding example of a natural process | UNESCO |
| N10 habitat | Site contains one of the most important natural habitats | UNESCO |
| size | The total area of a site in ha | UNESCO |
| Inscription date | Year of inscription on the WH list | UNESCO |
| transboundary | Whether the site crosses a national border | UNESCO |
| visitors | Annual number of visitors to a site | Own Survey |
| urban | Whether the site is either a city or located in one | UNESCO |
| archeology | Whether the site is an archeological site | UNESCO |
| danger | Whether the site has ever been put on UNESCO's danger list | UNESCO |
| conservation | Index from 0 to 5 in what condition the site is | Own Survey |
| distance city | Distance from site to next city in Km | Own Calucation |
| distance major city | Distance from site to next major city in Km | Own Calucation |
| distance coast | Distance from site to the coast in Km | Own Calucation |
| distance mountains | Distance from site to next mountains in Km | Own Calucation |
| mountain site | Site is located close to mountains | Own Calucation |
| coast site | Site is located close to the coast | Own Calucation |
| city10 | Number of cities within 10 Km of the WH site | Own Calucation |
| city50 | Number of cities within 50 Km of the WH site | Own Calucation |
| network30 | Number of cities within 10 Km of the WH site | Own Calucation |
| distance nature | Distance from site to the next natural attraction in Km | Own Calcuation |
| tourism10 | Number of tourists per country in 2010 | UNWTO database |
| tourism11 | Number of tourists per country in 2011 | UNWTO database |
| tourism12 | Number of tourists per country in 2012 | UNWTO datbase |
| tourism share | Tourism share in a country's exports 2010 | World Bank |
| state history | Score between 0 and 100, capturing how much of a country has been organized as a state for how long and by whom | Chanda and Putterman (2007) |
| HDI | Human Development Index, UNDP welfare estimator | UNDP database |
| GDP | Gross domestic product of a country | World Bank |
| federalistic state | Binary variable whether the state is federalistic (1) or not (0) | Democracy Cross National dataset (by Pippa Norris) |
| books export | Value of books ($) exported per country per year | |
| books import | Value of books ($) imported per country per year | |

| | |
|---|---|
| state history X culture site | Interaction term between a country's state history and its cultural WH sites. This captures a potentially distinct effect of a country's state history for cultural sites. |

cities with more than one million inhabitants or number of other WH sites within radii of different lengths.

Table 2 describes the explanatory variables used in the econometric framework.

Finally, we also collected data at the country level, to compare this to specifications in which country specific influences are controlled for with fixed effects. The additional country data comes from two sources: Teorell (2015), Chanda and Putterman (2007) and Norris (2009). Within these datasets, we use data on a country's history, tourism numbers, income, Human Development Index and cultural imports and exports.

When we finished our data collection (11/2011), there were about 900 WH sites and we collected data for 791 of them (from about 320,000 visitors). Since then, the number of WH sites has steadily grown to 1007 (2/2015). Furthermore, we have a full set of site specific explanatory variables for 453 WH sites and once we move from using country fixed effects to country specific data, we end up with a sample of 414 WH sites. To understand how this gradual sample size reduction affects the representativeness of our sample, table 1 shows characteristics for the full WH list as of 2/2015 and for the samples of 791 and 414, respectively. It can be seen that all our samples are representative both in the dimension region and in the dimension type.

Because the sites vary greatly in their tourism popularity, response rates vary equally for the feedback on WH visibility and site condition. In our sample of 453 sites, e.g., the mean response rate is 640 and the standard deviation is 1978. In section 7, we demonstrate this does not bias our estimates.

## III. Model

For our analytical framework, we assume that every location has inherent equity (unrelated to its WH status) in the minds of visitors, donors, agencies and the local public. Then, there is additional WH brand equity, which is the value premium that comes from the branding of the site as World Heritage. The overall equity of a site then is the sum of intrinsic equity and WH brand equity minus the costs of World Heritage status and its promotion and the costs of conserving the site.

As mentioned in the introduction, WH brand equity ($E_{it}^{WH}$) is a function of the value that the WH sites provide times how much people associate this value with the WH brand ($V_{jt} * B_{jt}$) as well as each site's characteristics ($S_{it}$) and its context ($C_{it}$):

$$E_{it}^{WH} = \sum V_{jt} * B_{jt} + S_{it} + C_{it} \tag{1}$$

This WH brand equity can be used by the associated sites to achieve their various goals, such as to attract more tourists or investors, or to convince local politicians to increase a site's protection.

Then, a location's equity ($E_{it}^{L}$) is a function of its inherent equity ($E_{it}^{I}$) plus the WH brand equity ($E_{it}^{WH}$) and, again, the site's characteristics ($S_{it}$) and its context ($C_{it}$) minus the costs to conserve the site $c(E_{it}^{I})$ and to be associated with the WH brand $c(E_{it}^{WH})$:

$$E_{it}^{L} = E_{it}^{I} + E_{it}^{WH} + S_{it} + C_{it} - c(E_{it}^{I}) + c(E_{it}^{WH}) \tag{2}$$

In other words, the WH sites create and benefit WH brand equity, but at highly different scales. Some sites (think of "Paris, Banks of the Seine") arguably create great value to any associated brand, while they do not benefit much from an additional association to the WH brand themselves. Other sites (think of a small, rather unknown site) might not have a large impact for an associated brand, while they might benefit greatly from a popular brand.

A straightforward way for the WH sites to contribute to WH brand equity is by raising their visitors' awareness of the association with the WH brand ($B_{it}$). This is why UNESCO encourages the sites to inform its visitors about the brand. However, one might suspect that without a clear requirement, only the sites that clearly and directly profit from their WH association do so. Hence, we estimate a model such as the following:

$$B_{it} = \beta_1 X_{it} + \beta_2 A_{it} + \beta_3 Z_{it} + \beta_4 Con_{it} + \varepsilon_{it} \qquad (3)$$

where $B_{it}$ is a measure for how much a location is branded as a WH site, $X_{it}$ is a vector of variables that affect how much an individual site benefits from its WH status, $A_{it}$ is a vector of variables that affect its ability to inform visitors about its WH status and $Z_{it}$ is a vector of covariates that also includes country fixed effects. Additionally, we probe the relationship between a site's branding and its conservation status, denoted $Con_{it}$.

Table 3 lists the variables that describe the benefits and costs of on-site branding.

If the sites are concerned with the collective brand equity and how much their contribution to it can help smaller locations, the coefficient $\beta_1$ should not be significant. In contrast, in this case we would expect $\beta_2$ to be highly significant. Regarding the relationship between branding and conservation status, there are theoretical reasons to expect either a positive or a negative relationship, depending on whether a stronger WH association increases unsustainable pressures or increases the interest and ability to protect the location (see i.e. Frey et al. (2011) and the literature discussed within).

To empirically identify equation (3), we employ a spatial econometric framework, to take into account the likely spatial dependence among observations.

**Table 3: Variables Describing Benefits and Cost of On-Site WH Branding**

| Benefit | Explanation | Cost | Explanation |
|---|---|---|---|
| network effect | -ability to free-ride reduces benefit of own marketing | transboundary | -Coordination and cooperation problems |
| distance from X | -benefit from global brand is e.g. larger in rural than urban regions | danger (listed on danger list in the past or currently) | -Indicator for management issues, e.g. insufficient resources |
| visitors | -benefit of WH is larger for sites with few visitors (for promotion) or many (for conservation) | size | -assumption that c.p. smaller sites are easier to manage |
| gdp | -benefit from global brand larger for developing countries | gdp | -Resources to manage WH site are more easily available in developed countries |
| kind of site | -interaction effect between brand and "product" | | |
| books import | -proxy for cultural interest | | |
| tourism | -benefit of global brand is c.p. greater the less tourism there is already | | |

Starting with a standard regression equation as shown in equation (3), the spatial dependence between observations can be specified in at least two ways (Ward and Gleditsch 2008, LeSage and Pace 2010):

First, we can allow the residual-term to be spatially correlated over observations (spatial error model), which can be interpreted as WH sites in a certain proximity sharing unobserved commonalities (i.e. similar visitors, infrastructure and institutional context). This model is specified as:

$$B_{it} = \beta_1 X_{it} + \beta_2 A_{it} + \beta_3 Z_{it} + \beta_4 Con_{it} + \lambda W\xi + \varepsilon_{it} \quad (4)$$

where W is a spatial weights matrix (measuring the proximity of the locations), $\xi$ denotes the spatially correlated residuals of close-by sites and $\lambda$ is a coefficient to be estimated.

Alternatively, spatial dependence can also be captured by allowing the on-site WH visibility of sites to directly influence the on-site WH visibility of close-by sites (spatial lag model). This could arise i.e. if site managers in a region influence each other (i.e. freeriding on the promotion efforts of other sites) or cooperate (i.e. to establish a regional tourism profile). This second model is specified as

$$B_{it} = \beta_1 X_{it} + \beta_2 A_{it} + \beta_3 Z_{it} + \beta_4 Con_{it} + \rho W B_{jt} + \varepsilon_{it} \quad (5)$$

where $B_{jt}$ denotes the WH on-site visibility of close-by sites and $\rho$ is a coefficient to be estimated.

Both specifications (4) and (5) can readily be estimated in Stata (Pisati 2001) or Matlab (LeSage 1999).

## IV. Descriptive Findings

Table 4 shows the World Heritage sites that are most and least branded as such. Figure 1 shows the geographically aggregated visibility in countries and world-regions. The highest average visibility is found in the Asia-Pacific region, the lowest in the Middle East. Individually analyzed, "L'Anse aux Meadows National Historic Site", an $11^{th}$-century Viking settlement in Canada has the highest reported visibility, followed by the "Shark Bay", a nature site in Western Australia that features high biodiversity. The "Shark Bay" is followed by "Ironbridge Gorge", a complex from the industrial revolution in the UK, the "Wet Tropics of Queensland", an Australian rainforest site, featuring a particularly rich fauna and flora including endangered species, and the "Heart of Neolithic Orkney", a group of monuments in the UK that give a graphic depiction of life in this region 5000 years ago. Furthermore, high visibilities are found at the "Church Village of Gammelstad" in Sweden, the "Head-Smashed-In Buffalo Jump" in Canada and the "Sun Temple" at Konarak, India. The lowest WH visibility is reported for the "Selous Game Reserve" in Tanzania, which is a large, undisturbed park, hosting a large variety of mammals and other animals and

**Table 4: World Heritage sites with especially high or low brand visibility**

| Highest WH Visibility | | | Lowest WH Visibility | | |
|---|---|---|---|---|---|
| Name | Country | Visibility | Name | Country | Visibility |
| L'Anse aux Meadows | Canada | 4.78 | Selous Game Res. | Tanzania | 2.16 |
| Shark Bay | Australia | 4.74 | Paris, Banks of the S. | France | 2.24 |
| Ironbridge Gorge | UK | 4.58 | Medina Marrakesh | Morocco | 2.28 |
| Wet Tropics of Queensland | Australia | 4.5 | Medina Sousse | Tunisia | 2.31 |
| Heart of Neolithic Orkney | UK | 4.46 | Medina Tunis | Tunisia | 2.32 |
| Gammelstad | Sweden | 4.46 | Talamanca /-La Amistad | Panama, CR | 2.33 |
| Head-Smashed-In Buffalo J. | Canada | 4.43 | Kilimanjaro National Park | Tanzania | 2.35 |
| Historic Centre of Telc | Czech R. | 4.41 | Historic Centre of Naples | Italy | 2.37 |
| Lord Howe Island Group | Australia | 4.40 | Canal ring Amsterdam | Netherlands | 2.37 |
| Sun Temple, Konarak | India | 4.40 | Medina Essaouira | Morocco | 2.40 |

"Paris, Banks of the Seine", which covers large parts of Paris' city center, which is inscribed for its broad influence on 19th and 20th century world-wide town planning. Furthermore, the Medinas of Marrakesh, Sousse, Tunis and Essaouira, which are located in Morocco and Tunisia (inscribed for their architecture, art and historical significance) all appear in the list of very low WH visibility - as do the "Talamanca /-La Amistad Reserve and Park" in Panama and Costa Rica (an area of tropical rainforest where species from North and South America mix), the "Historic Center of Naples" in Italy and the "Seventeenth-century canal ring area of Amsterdam" in the Netherlands. The best conservation status of all WH sites is reported for "Roskilde Cathedral" in Denmark and "L'Anse aux Meadows" in Canada.

It can also be seen that the best conserved sites are exclusively found in Europe and North America. The worse conservation status is reported for "Kasbah of Algiers" in Algeria and the "Fort and Shalamar Gardens" in Pakistan. Most countries with conservation issues are in developing regions.

**Table 5: World Heritage sites in especially good or bad condition**

| Best Reported Condition | | | Worst Reported Condition | | |
|---|---|---|---|---|---|
| Name | Country | Condition | Name | Country | Condition |
| Roskilde Cathedral | Denmark | 4.8 | Kasbah of Algiers | Algeria | 2.3 |
| L'Anse aux Meadows | Canada | 4.8 | Fort and Gardens Lahore | Pakistan | 2.8 |
| Lord Howe Island Group | Australia | 4.7 | Island of Saint-Louis | Senegal | 2.9 |
| Hal Saflieni Hypogeum | Malta | 4.7 | Hattusha | Turkey | 3.0 |
| Places in Nancy | France | 4.7 | Harar Jugol | Ethiopia | 3.0 |
| University of Virginia | USA | 4.7 | Historic Centre of Naples | Italy | 3.1 |
| Head-Smashed-In Buffalo J. | Canada | 4.7 | Chola Temples | India | 3.1 |
| Chaco Culture | USA | 4.7 | Ciudad Univ. de Caracas | Venezuela | 3.1 |
| Heart of Neolithic Orkney | UK | 4.7 | Medina of Touan | Morocco | 3.1 |
| Garajonay National Park | Spain | 4.7 | Bahla Fort | Oman | 3.1 |

**Figure 1: The On-Site Visibility of the World Heritage Brand in Countries and Regions**


Map: Darker areas indicate higher WH visibility while lighter areas indicate lower WH visibility. White areas denote missing data. Boxplots: The boxes represent 50% of the variation in WH visibility, the line in the middle denotes the median. The lines outside the boxes represent 75% of the variation and the dots show outliers. NA = North America.

## V. Analytical Findings

Table 6 presents the results of four specifications that aim to explain which sites are better branded as WH sites than others. Two specifications are spatial error models and two specifications are spatial lag models. The first two specifications use country fixed effects to control for country specific influences (such as kind of tourists, infrastructure or culture). The latter two specifications use only continental fixed effects but include various country specific variables.

**Table 6: Spatial Regression Results: Who promotes the World Heritage Brand?**

| | Spatial Err. 1 | | Spatial Lag 1 | | Spatial Err. 2 | | Spatial Lag 2 | |
|---|---|---|---|---|---|---|---|---|
| | Coef. | S.E. | Coef. | S.E. | Coef. | S.E. | Coef. | S.E. |
| distance nature | -0.060*** | (0.022) | -0.060*** | (0.022) | -0.063*** | (0.021) | -0.0707*** | (0.021) |
| distance city | 0.086*** | (0.021) | 0.086*** | (0.021) | 0.078*** | (0.021) | 0.0853*** | (0.021) |
| network effect | -0.075*** | (0.017) | -0.075*** | (0.017) | -0.082*** | (0.018) | | |
| vsitors | -0.601*** | (0.151) | -0.601*** | (0.151) | -0.410** | (0.167) | -0.354** | (0.170) |
| visitors squared | 0.580*** | (0.151) | 0.580*** | (0.151) | 0.395** | (0.167) | 0.335** | (0.169) |
| conservation | 0.160*** | (0.022) | 0.160*** | (0.022) | 0.195*** | (0.022) | 0.195*** | (0.022) |
| archeology | 0.042*** | (0.016) | 0.042*** | (0.016) | 0.017 | (0.017) | 0.0163 | (0.018) |
| urban | -0.058*** | (0.017) | -0.058*** | (0.017) | -0.082*** | (0.019) | -0.0934*** | (0.019) |
| culture site | 0.040 | (0.034) | 0.040 | (0.034) | 0.072** | (0.036) | 0.062 | (0.051) |
| nature site | -0.012 | (0.033) | -0.012 | (0.033) | 0.03 | (0.038) | -0.007 | (0.041) |
| transboundary | | | | | | | -0.011 | (0.018) |
| size | | | | | | | -0.007 | (0.020) |
| Inscription date | | | | | | | -0.019 | (0.020) |
| danger now | | | | | | | -0.006 | (0.020) |
| danger | | | | | | | 0.002 | (0.020) |
| state history X culture site | | | | | -0.030** | (.015) | | |
| C1 masterpiece | | | | | | | -0.127*** | (0.040) |
| C2 interchange | | | | | | | -0.046 | (0.039) |
| C3 past tradition | | | | | | | 0.065 | (0.040) |
| C4 history | | | | | | | 0.031 | (0.044) |
| C5 interaction | | | | | | | 0.016 | (0.057) |
| C6 living trad. | | | | | | | -0.049 | (0.042) |
| N7 phenomenon | | | | | | | 0.020 | (0.102) |
| N8 history | | | | | | | 0.077 | (0.118) |
| N9 processes | | | | | | | -0.110 | (0.101) |
| N10 habitat | | | | | | | 0.072 | (0.109) |
| books Export | | | | | | | -0.003 | (0.044) |
| books Import | | | | | | | 0.263** | (0.102) |
| tourism 2012 | | | | | -0.084*** | (0.021) | -0.156*** | (0.031) |
| federalism | | | | | 0.017 | (0.017) | 0.00834 | (0.020) |
| state history | | | | | -0.014 | (0.021) | 0.00893 | (0.023) |
| HDI | | | | | 0.019 | (0.035) | | |
| GDP | | | | | | | -0.0634** | (0.028) |
| Country FE | Yes | | Yes | | No | | No | |
| Continent FE | No | | No | | Yes | | Yes | |
| pseudo r2 | 0.73 | | 0.73 | | 0.49 | | 0.46 | |
| sigma | 0.29 | | 0.29 | | 0.35 | | 0.34 | |
| observations | 453 | | 453 | | 414 | | 414 | |

Significance levels are 10% = *, 5% = ** and 1% = ***

This reduces the pseudo $R^2$ from about 70 to about 50 and the squared correlation from 60 to 45. Nevertheless, the estimated effect of variables included in all specifications is remarkably constant.

It can be seen that sites are better branded if they are closer to natural attractions (distance nature) and further away from cities (distance city). Also, sites in better condition (conservation) are better branded as are those with few or many visitors (visitors and visitors squared), in contrast to sites with intermediate visitor numbers. The type of site is not really an important factor (culture site, nature site), as in three out of four specifications the distinction between cultural, natural and mixed sites does not matter and from all inscription criteria (C1 – N10), only one is estimated to be significant (C1, Site represents a masterpiece of human creative genius). As an exception to this, urban sites (urban) are worse branded than other sites, while archeological (archeology) sites have a tendency to be better branded (which could once more reflect the urban-rural distinction also shown by the effects of distance to next city and natural attraction). Also, for cultural sites the history of the country matters (state history): Cultural sites located in a country with a longer state history (which hence had more time to accumulate cultural heritage) are branded less (state history X culture site).

Variables that capture how difficult the branding might be (transboundary sites, being on UNESCO's danger list, a site's size) are not significant and neither is the date of inscription. Regionally, there is a strong positive effect for sites in the Asia-Pacific region, an intermediate positive effect for Europe and North-America and a weak positive effect for Latin-America, which is relative to Arab States and Africa.

Regarding country specific variables, the number of international tourist arrivals has a negative impact on a site's WH branding (because the incentive is smaller in countries with already popular tourist destinations), while its value of annual book imports is positively correlated (reflecting a higher general cultural interest in these countries) with WH branding. Countries with higher GDP are slightly less branded than countries with a lower GDP.

We also tested estimating individual models for different types of WH sites (i.e. cultural or natural WH sites) but there are no heterogeneous effects to be found.

## VI. Robustness Checks

In this section ,we estimate 3 more specifications, to investigate the robustness of our previous results.

Above, we find that WH sites in countries that attract more international tourists are less promoted as WH sites than WH sites in countries that attract less international tourists. Our measure is the number of international tourist arrivals per year (table 6). To test what we really capture, we can instead, or in addition, also use the tourism share in a country's exports as a measure for a country's tourism specialization (Arezki et al. 2009).

To test whether different response rates at the various WH sites affects our estimates, we can directly include this variable in our model. Just as we did with the visitor numbers, we allow the response rate to have a non-linear effect by including an additional squared term. Table 7 presents 3 specifications: 2 spatial error models and 1 spatial lag model.

Regarding the tourism specialization, it can be seen that this is not a relevant dimension for the on-site promotion of the WH brand. International tourist arrivals remain highly significant when included (specification 1) and tourism specialization is never significant, independent from the inclusion of international tourist arrivals. Regarding the effect of the response rate it can be seen that it merely reflects the visitation pattern at the sites but with less measurement error (see section 2). It can be seen that our estimations above are unbiased by the different response rates at the sites and for ease of interpretation, it seems preferable to exclude the response rate from the model, as we cannot see the effect of the visitation pattern when it is included, because of differential measurement error.

**Table 7: Robustness Checks**

| | Spatial Err. 1 | | Spatial Lag 1 | | Spatial Err. 2 | |
|---|---|---|---|---|---|---|
| | Coef. | S.E. | Coef. | S.E. | Coef. | S.E. |
| distance nature | -0.0559*** | (0.0208) | -0.0633*** | (0.0214) | -0.0546** | (0.0213) |
| distance city | 0.0710*** | (0.0209) | 0.0753*** | (0.0217) | 0.0653*** | (0.0217) |
| network effect | -0.0656*** | (0.0182) | | | -0.0593*** | (0.0185) |
| visitors | -0.132 | (0.179) | -0.0380 | (0.184) | -0.108 | (0.183) |
| visitors squared | 0.122 | (0.178) | 0.0244 | (0.183) | 0.0951 | (0.182) |
| culture site | 0.0661* | (0.0357) | 0.0681 | (0.0465) | 0.0739 | (0.0460) |
| nature site | 0.00574 | (0.0351) | -0.00119 | (0.0413) | 0.00141 | (0.0408) |
| conservation | 0.191*** | (0.0204) | 0.203*** | (0.0220) | 0.200*** | (0.0217) |
| archeology | 0.0142 | (0.0173) | 0.00886 | (0.0184) | 0.0115 | (0.0182) |
| urban | -0.0773*** | (0.0187) | -0.0774*** | (0.0195) | -0.0756*** | (0.0192) |
| tourism 2012 | -0.0810*** | (0.0189) | | | | |
| tourism_share | 0.0139 | (0.0305) | 0.0150 | (0.0322) | 0.00710 | (0.0319) |
| GDP | | | -0.0604** | (0.0278) | -0.0557** | (0.0275) |
| danger now | | | -0.0117 | (0.0198) | -0.0105 | (0.0196) |
| danger | | | 0.00992 | (0.0198) | 0.0159 | (0.0197) |
| books Export | | | 0.0467 | (0.0415) | 0.0383 | (0.0410) |
| books Import | | | -0.0421 | (0.0392) | -0.0362 | (0.0388) |
| responses | -0.165*** | (0.0379) | -0.172*** | (0.0384) | -0.153*** | (0.0384) |
| responses sq | 0.113*** | (0.0355) | 0.109*** | (0.0362) | 0.106*** | (0.0358) |
| Kind of site | Yes | | Yes | | Yes | |
| Continent FE | Yes | | Yes | | Yes | |
| R-sq | 0.46 | | 0.44 | | 0.46 | |
| AIC | 323.8 | | 358.0 | | 349.9 | |
| BIC | 408.3 | | 490.9 | | 486.7 | |
| N | 414 | | 414 | | 414 | |

Significance levels are 10% = *, 5% = ** and 1% = ***

## VII. Discussion

The estimates presented in table 6 suggest that destinations that benefit more from the WH brand contribute more to its equity and sites that benefit less also contribute less. Thus, we find a complementarity in the production of private and collective benefits (Cornes and Sandler 1984).

In our case, however, on-site promotion of the WH brand is not a direct means of producing a private benefit as visitors have already decided to visit when they are approached. Rather, the sites that benefit from their WH branding are more willing to contribute to the collective brand, so they deliberately work towards strengthening the brand when they believe to benefit from it.

We find the sites' location to be important and in particular, in which country it is and how close to the next city. Clearly, the brand is differently popular in different countries (high e.g. in Asian countries and low in the Middle East); and sites that are close or even within a city perhaps find it easier to attract visitors, ceteris paribus, because the costs of visiting are low (both in terms of money and time), while rather remote sites, need to convince visitors to come and for this, the WH label arguably helps (in two ways: first, the WH list creates awareness for rather less known locations and second, it functions as a quality mark, certifying how special and valuable the location is). In contrast, proximity to a natural attraction encourages WH branding. Perhaps, this is because natural attractions are complementary to cultural attractions, such as a WH site, so that the benefit of promoting a location as a WH site is greater in combination with the right environmental background. Somewhat similar, archeological sites are found to be better branded, which we interpret as evidence for some sites' themes having more potential to be market as WH than others. In this regard, it is also important how much competition there is from other brands. While for archeology, WH is perhaps an easy choice, for a city or a national park, there are perhaps several attractive alternatives for how to promote these locations.

The finding of a negative network effect is in line with the finding of a negative effect of a country's tourism popularity. Basically, this suggests that sites use the WH brand more actively if they cannot profit from the general or specifically cultural tourism popularity of their region.

The relationship between visitor numbers at the sites and their branding is U-shaped. Is this relationship causal? That is, are sites that attract more visitors better branded, or are sites that are better branded more visited? Since we analyze on-site branding, this is unlikely to attract many more visitors (as only actual visitors get to see the branding on-site). Hence, it is plausible that it is the visitor number that affects the branding. The reason why we can nevertheless explain the WH branding of a site partially with its economic incentive to do so is the following: The on-site branding does not increase revenues but it is sites that profit economically from the WH brand that value it higher. Thus, sites that are especially prone to attract more or higher spending tourists due to their WH listing contribute the most to increasing WH brand equity because they have more appreciation of the brand.

Importantly, the benefit of the WH brand is not limited market benefits. First of all, we find that sites with especially many tourists are among the better branded. It is unlikely that these sites try to attract still more tourists but more likely, they use the WH brand to get support to protect the site. It is well known that tourism pressure can endanger the integrity of a WH site, so perhaps, the WH brand is used to persuade decision-makers and the public to ensure a sustainable use.

In the same direction, we find a positive relationship between a site's conservation status and its branding. This relationship is unlikely to be causal but more likely reflects a virtuous cycle, in which a better site condition creates economic value, which encourages commitment to the WH mission which in turn supports the preservation of the site.

All of the above suggests that it is the site specific benefits that determine branding. But what about the costs of branding? In short, we do not find evidence that constraints significantly matter. Naturally, some sites should have it easier to brand themselves (small, compact sites, that are within a single country and have been inscribed for many years, etcetera), but none of the potential variables are significant.

A variable that could affect the ability of a site to invest into branding and also its economic incentive to do so is a country's GDP. As Arezki et al. (2009) show, tourism specialization can be a viable development strategy. Hence, the effect of GDP could be positive (if a higher GDP gives a country more resources to invest into the branding of its WH sites) or negative (if less developed countries have a higher incentive to use the WH brand for development). As shown in table 5, the incentive trumps the financial constraint.

In summary, we find empirical evidence that it is predominantly the economic benefit of the sites themselves that affects their association with the WH brand. But this effect is indirect, as the sites that benefit more promote the brand on-site without directly increasing their benefits by doing so.

We think the logic of our model can even be turned around. Instead of using the incentive to use the WH brand to explain how much sites are branded, perhaps one could use our measure of on-site branding to predict which sites benefit more from their association with the WH brand. This could be done in future research to explore whether it is possible to help more locations to benefit from their WH inscription, which seems particularly worthwhile given our finding of a strong positive correlation between WH association and conservation status.

## 8. Conclusion

We quantitatively investigate the extent to which WH sites are clearly branded as such and thereby promote the collective brand. To this end, we collected responses of almost 320,000 visitors at 791 WH sites on how aware they were of the WH logo, pamphlets, brochures or other information material concerning the WH brand.

We then collected data on explanatory variables and estimated spatial econometric models on a subsample of 451 and 414 WH sites, respectively, to identify the mechanisms behind the choice to clearly brand a location as WH site.

We find site specific incentives to explain this decision well.

This suggests that our big-data measure of WH branding might also be used as an indicator for which sites benefit from WH inscription. Generally, we find that sites that are further away from a city but closer to the coast or mountains are more likely to be branded and that the brand is more popular in Asia, followed by Europe and North America, while relatively unknown in the Middle

East and Africa. Furthermore, the effect of visitor numbers follows a U-shape, while cultural sites in culturally richer countries are less branded as are sites in economically richer countries and those with already more tourism popularity.

This suggests that site management is to a large extent driven by market values and also, but less, by non-market and cultural values.

The policy implications of our research are as follows: The inconsistent WH brand visibility throughout the WH network arises in part due to the absence of a unifying communications framework under the WH Convention. Though all sites are said to be humanity's common heritage, there appears to be nothing common in the way individual sites perceive and/or implement their obligations regarding the telling of the WH story. The only formal reference regarding the need to include a clear WH logo and messaging at WH sites is limited to mere suggestions on the design and location of a plaque – there are no formal requirements in the text of the WH Convention nor in that of its Operational Guidelines for systematic corporate messaging (UNESCO 1972, 2013). Hence, the obligation to contribute to collective brand equity that is usually imposed upon members of a collective brand is absent. Communications requirements imposed upon WH sites are only suggested. At the same time, it is clear that WH stakeholders, both at the site level and at the global level, should do what they can to encourage a more systematic and effective WH brand awareness building. As a proposition, the WH Committee, charged with managing the implementation of the WH Convention, may wish to consider modifying the Convention's Operational Guidelines so that communicating the WH message be made a more explicit and mandatory component of joining the WH family, for example.  Similarly, proponents of new WH sites should be requested to provide a clear WH communication plan at the site, against which it can be evaluated if inscribed onto the WH list.

Independent from making the WH brand association mandatory, our second recommendation is to further explore the topic which, where and how locations benefit from their WH status. There is a lively debate over which sites attract significantly more visitors because of their WH and how WH listing affects site management. Once the determinants are identified and patterns emerge, it is possible to better support sites that currently benefit less than average and to share best practices around the world.

# CONCLUSION: CULTURE AND RURAL DEVELOPMENT

In chapter 1 and 2, we have seen how cultural traits can be crucial determinants of economic behaviors and performance. The smallholder pineapple farmers of Ghana invest in domains for which they hold high degrees of self-efficacy and they avoid domains for which they hold low degrees of self-efficacy. The domain specific self-efficacy in turn is found to be the outcome of historic experiences. Colonial experiences with cocoa cooperatives and Christian missionary schools continue to shape what the farmers expect and what they can accomplish with their communities. This is especially important for the performance of contract farming (chapter 1). Pre-colonial production systems similarly continue to affect the cultural beliefs of the farmers. Specifically, where historic production systems sufficiently rewarded investments, the farmers are currently more willing to invest, whereas farmers in other regions of Ghana are more reluctant. This is found to have significant income implications (chapter 2).

In chapter 3 we also see that farmers with high self-efficacy respond quite different to insufficient rainfall than farmers with low self-efficacy. Whereas the former respond by increasing investments into technology, the latter do not respond at all.

However, history and culture are not determining incomes. In chapter 4, education and non-historically determined social capital are found to help farmers to achieve higher incomes than predicted from their ancestral heritage (the experience of the trans-Atlantic slave trade and pre-colonial production systems that discouraged investments).

Furthermore, when the adoption of innovations is mostly constrained by information, chapter 5 presents empirical evidence that the provision of trainings can be highly beneficial for the farmers – but it also shows that trainings in less information constrained innovations is inefficient.

In the final chapter, two different aspects of culture are investigated. First of all, this is built culture, in the form of UNESCO's World Heritage sites. Secondly, it is found that intangible culture strongly affects the preference for the World Heritage program. The research question in chapter 6 is what determines that some World Heritage sites promote themselves as such, whereas others do not. This is an important question as the brand equity of the World Heritage program is mostly determined by its affiliated sites and especially its benefit to less famous locations is a function of this brand equity. Put simply, super star sites make the World Heritage brand more valuable – if they are promoted as World Heritage – and this value can be used to achieve development and conservation goals – especially at sites that are less popular amongst tourists, donors and policy makers. The explanation why some sites promote themselves as World Heritage sites include mostly tourism and location specific factors. Thus, standard microeconomics is a good explanation for the management but culture is clearly also a part of the explanation. Especially in Asia, the brand equity of the World Heritage brand is high, whereas it is almost non-existent in the Middle East. An interesting detail is the finding that countries with longer history as states have cultural heritage sites that are branded less as such, suggesting that if a country is well established as having a rich cultural history, there is little incentive to use a global brand as promotion argument. In contrast, it is these countries that have the greatest potential to contribute to World Heritage brand equity.

The main ambition of this thesis was an advancement of our understanding what culture actually is in an economic context, and how it affects rural development. A special focus is put on self-efficacy. This concept is commonly viewed as a personality trait, and thus categorized into psychology. However, the research presented in the chapters above identifies historical roots of self-efficacy – roots that are far deeper than individual lifes. What can be found is cultural evolution, describing the process of economic agents inheriting heuristics (simple decision rules to cope in complex environments), augmenting and changing them and passing them on to younger generations. This process is generally superior over individuals having to learn everything on their own but it comes at the cost that behavior is sometimes better adapted to past circumstances (of earlier generations) than to the current ones. Indeed, it can be found that especially self-efficacy is often too low, meaning that many economic agents could be made better off, if they would be more convinced to have the ability to achieve success in important domains. As covered in chapter 1, contract farming could be far more efficient in Ghana, if the involved stakeholders would be more convinced that they can benefit from it (an almost self-full-fulfilling expectation). Similarly, many farmers could significantly increase their income if they would increase their investments, but they are constraint by low investment self-efficacy (chapter 2) and as shown in chapter 3, the negative impact of low rainfall could be mitigated if the farmers were more convinced that their adaptation behavior is effective.

As discussed in the introduction, there is a spectrum of determinants for economic development. Self-efficacy is fundamental, influencing many determinants at higher levels. By motivating economic decision makers to pursue more ambitious goals and to work harder and more resilient on their achievement, environmentally induced differences in self-efficacy are good candidates for a theory about why some regions in the world have better institutions, more human, social and physical capital, more entrepreneurship and higher levels of savings and investment.

In the chapters 4 and 5, factors are explored that could compensate for low self-efficacy. Such factors are human capital (both practical and formal) and social capital (especially caused by frequent social interaction). As argued above, human and social capital are outcomes of self-efficacy, so it seems strange at first to argue that these are also factors that can compensate for low self-efficacy. The solution to this puzzle are positive feedback effects. Increasing self-efficacy is likely to increase human and social capital but externally increasing those variables also increases self-efficacy. This is also important to understand because it suggests virtuous and vicious cycles. Thus, small initial differences can put regions and individuals onto entirely different development paths.

Investigating economic culture also poses interesting methodological challenges. First of all, it is not trivial to accurately and precisely measure cultural traits. Secondly, culture is almost always endogenous. Arguably, this is one reason, why econometric studies have neglected culture for a long time until recently.

E.g. in the chapters 1 to 3, we want to identify the causal effect of different degrees of self-efficacy on (1) the performance of contract farming, (2) annual incomes, and (3) the response to decreasing rainfall. The challenge is that self-efficacy is obviously endogenous. For this reason, much research relies on randomized control trials, which has, however, serious shortcomings. First of all, self-efficacy can be characterized by its magnitude ("effect-size"), strength (how resilient it is) and generality (whether it only applies to a small domain, or includes a spectrum of domains). In experimental research, self-efficacy is usually manipulated by the provision of bogus information (e.g. manipulated feedback to a task). This changes the self-efficacy of the treatment group but only creates weak self-efficacy, because the development of strong self-efficacy requires sufficient "mastery experiences" of the individual, which needs time. Telling a group of people that they are well equipped to be successful in a domain increases their aspiration and motivation but if they fail in the domain, their self-efficacy quickly adapts. In contrast, a person with strong self-efficacy is not much affected by failure and adversity. Thus, experimental studies on self-efficacy only investigate a small aspect of self-efficacy – usually increased aspirations or initial motivation. A second shortcoming of experimental studies is that because degrees of self-efficacy are manipulated by the researcher, little is learned about the effect of actual differences in self-efficacy. If differences in self-efficacy are larger or smaller in reality than what is created in the experiment, its effect might be under- or overestimated. Another consequence of experimental manipulation is that the actual sources of self-efficacy are not identified. In the experiment, the source is the researcher but it is also of interest where real life differences in self-efficacy come from. For these reasons, non-experimental data might be sometimes preferred – not as a substitute but as a complement to the experimental studies. However, the challenge then is to exogenize self-efficacy to identify its causal effect, which can be achieved by using instrumental variables in an econometric framework. Feasible instruments are suggested by self-efficacy theory and the theory of cultural evolution. The former states that self-efficacy is socially learned and the latter states that social learning does not only happen horizontally (within a generation of social peers) but also vertically (between generations). Thus, the search for historic circumstances that affected the evolution of self-efficacy (but that do not otherwise affect the outcome of interest) is clearly tedious but with a high return on investment. Such historic circumstances can provide exogenous variation to the currently measured self-efficacy of individuals. However, historic instruments require much testing, because the long time difference between the instrument and the measurement of the endogenous variable potentially allows many alternative causal channels to interfere. The approach taken in the chapters above is to perform falsification tests, e.g. to use the chosen instruments to explain an alternative causal channel.

To connect historic instruments with current observations, two strategies are employed. First, using data about the historic locations of African ethnicities by George P. Murdock, historic variables can be linked with current observations through the reported ethnicity of the individuals. The second approach is to use a geographic information system (GIS) to link historic events to individuals based on their location. Theoretically, one approach is as good as the other – in our case, geographic linking produces much more variation than ethnic linking.

Based on self-efficacy theory, a feasible instrument for self-efficacy in a specific domain must be a historical event in the same or a sufficiently similar domain. As an example, in chapter 1 the domain is (pineapple) contract farming. The instrument for self-efficacy in this domain is the historical performance of colonial (cocoa) cooperatives. An advantage of the colonial cocoa cooperatives is that they were established everywhere in Ghana (in the 1930s), but not by the farmers but by the colonial government (the British). A problem, however, is that this instrument is not exogenous either. Regions with a higher average degree of self-efficacy might have been more successful with the cocoa cooperatives during colonial times and these regions are currently more successful with pineapple contract farming. In this case, we would estimate the causal effect of persistent differences in self-efficacy but not of self-efficacy changed during Ghana's colonial era. This matters because policy recommendations would be different. Finding persistent self-efficacy would be consistent even with a genetic explanation, or persistent environmental influences, with little room for policies. Finding, however, that colonial experiences affected the evolution of self-efficacy suggests the possibility to affect self-efficacy with policies and new experiences and also improves the reliability of our identification strategy (otherwise it would be difficult to disentangle genetic, environmental, and cultural effects.). To exogenize our instrument, we use a set of "secondary instrument" – variables that affected the performance of the colonial cocoa cooperatives but not otherwise the current performance of the pineapple farming contracts. Such instruments include especially the historic rainfall on the cocoa farms, their suitability to grow cocoa, the distance to the colonial railway. These instruments do not affect the current performance of pineapple farming contracts, conditional on controlling for the rainfall on the pineapple farms, their suitability to grow pineapple, and by acknowledging that whereas cocoa was transported with the railway to the coast, the pineapples are transported with trucks to processing companies or directly sold locally.

Thus, it is possible to exogenize the colonial performance of cocoa cooperatives with historic environmental and infrastructure variables and then use it to exogenize the self-efficacy of Ghana's pineapple farmers.

These instruments do not only help to identify the causal effect of self-efficacy on contract farming performance, they are also of individual interest. Finding that current degrees of self-efficacy are caused by colonial experiences should caution current development programs and policies, considering their potential long-term impact. Program or policy failures might have a long shadow.

A similar model is also developed in chapter 2, where the question is whether self-efficacy differences produce income differences. To motivate the econometric model, an agent based model is developed, from which all hypotheses are derived. The first hypothesis is that income differences are mainly caused by investment differences, which are in turn caused by investment self-efficacy differences. The second hypothesis is that historic environments differed in their incentive to invest and thus, the farmers in different regions made different investment experiences, which caused different degrees of investment self-efficacy. To empirically validate the agent based model, a historic variable is needed that affect the incentive to invest. Historic production systems are such a variable. Whereas cereals incentivize agricultural investments, roots, tubers and tree crops do so much less. Thus, historic dependency on different agricultural subsistence systems is a good explanation for current differences in investment self-efficacy. However, this instrument is potentially endogenous, if societies with higher self-efficacy self-selected into cereal farming, and societies with lower self-efficacy self-selected into other production systems. To get rid of this potential selection bias, data from the FAO on regional differences in the suitability to grow different crops is exploited. The comparative advantage to grow a given crop indeed explains well why this crop was grown in a given region and this in turn explains differences in self-efficacy. In a second step, mediation analysis is used to show that the causal chain from historic production systems over differences in self-efficacy to income differences is consistent with the empirical data.

This demonstrates that self-efficacy can be robustly analyzed without experimental data but it requires complicated research designs and advanced econometric methods, which are naturally more sensitive to misspecifications than simple designs analyzed with standard methods.

An issue that has been omitted, until now, is how to capture somebody's self-efficacy in the first place. Asking individuals about their self-efficacy directly is not reliable, even if the respondents try to be honest. Instead, in the first chapter, respondents are openly asked about past income determinants, and the answers are coded into low, medium, and high self-efficacy. Identification is ensured by instrumenting this variable with the colonial cooperatives' experiences, so that only the part of the reported income determinants that is explained by this instrument is used as explanatory variable in the reduced form.

In chapter two, a battery of questions is used to create a factor variable, that summarizes different aspects of self-efficacy, including the open question from chapter 1, and also questions about one's relationship to the environment, planning horizon, investment perceptions, and perceptions of one's ability. Again, the final part of the identification strategy is to only use the part of this factor variable that is explained by the historic instrument in the reduced form.

In chapter 4, it is investigated what is needed for farmers to achieve a higher income than predicted given the region's history - specifically the trans-Atlantic slave trade and the above discussed historic production systems. These factors to overcome historic constraints are also partially endogenous, especially social capital, which is found to be important. Social capital is well known to be a crucial ingredient for economic development but it is an outcome of the characteristics and behaviors of the individuals. To instrument social capital, we cannot use a historic instrument, because we are exactly interested in non-historic factors. Such a non-historic factor is the frequency of social events attendance. Social interactions increase social capital but do not have other economic relevance. To capture the complex concept of social capital in the first place, a factor variable is estimated from a generalized trust question, the reported number of people who would lend money to the farmer and her attendance of social events. Thus, the attendance of social events is both an input to the factor variable and an instrument to exogenize it.

Also in chapter 5 do we run into endogeneity. The interest is on the effect of training on the adoption of innovations. This requires that we also identify the effect of peer-learning but both training and peer learning are endogenous. Participation in training could theoretically even follow the adoption of an innovation, a problem that is partially mitigated by our use of panel data, but the data only has observations on an annual basis, so an innovation could have been adopted in spring and then training could have been attended in the summer. Homogenous behavior within networks could indicate peer learning but it needs to be disentangled from individual responses to common shocks and the effect of shared individual characteristics. The instruments for both endogenous variables (training and networks) are the lagged values from indirect neighbors. The reason is that e.g. the network variable is defined as the average diffusion of an innovation in a community. If this average diffusion is caused by a factor that also affects the individual adoption decision, then we have selection bias. The average diffusion of the innovation in an adjacent community, however, does not capture the omitted factor that causes this selection bias – by construction. The identification of training is helped by the understanding that these are provided in one community and then in the next, and that participation approximates 100% within communities. Knowing how many farmers have been trained in one community is thus a strong predictor of how many farmers are going to be trained in the next.

In chapter 6, the question is how much the World Heritage sites brand themselves as such. As the World Heritage sites are spread all over the globe, data collection could seem prohibitively expensive. However, partnering with a popular online travel platform, visitors to the World Heritage sites could be inexpensively approached through the website and asked a few questions about their perceptions at the sites. This approach generated a very large data set of 300.000 observations. To robustly explain the determinants of World Heritage branding, additional data was collected about the sites, their host countries and their wider regions. To take into account spatial effects, the econometric framework allows both for unobserved but common arguments in the residual term (e.g. a shared, local institutional context) and for network effects (e.g. freeriding of a site on the marketing efforts of its neighbors).

In conclusion, it is found that culture is important for economic development and it can be rigorously analyzed with existing quantitative methods. This is encouraging, as a main criticism for economics has been the ignorance of social interactions and culture, of which the former has already found widespread application in recent economic models and the latter is just starting to follow.

Table 1 summarizes the main findings of this thesis.

## Table 1. Contributions of the previous chapters

| | Topic | Question | Hypothesis | What is new |
|---|---|---|---|---|
| 1 | Contract Farming | Why are some farmers profiting more from contract farming than others in Ghana? | For historical reasons, some farmer have higher self-efficacy and higher social capital | Whereas studies usually analyzed whether contract farming is profitable for the farmers - here the question is for whom it is profitable. Furthermore, the psychological concept of Self-Efficacy is introduced into agricultural economics and it is shown that it has historical roots, which makes it a cultural trait. |
| 2 | Farm Incomes | Why do some farmers in Ghana have higher incomes than others? | For historical reasons, some farmers have higher self-efficacy regarding investments, which is why they invest more and have higher incomes | Methodologically, agent based modelling and econometrics are combined. Furthermore, it is investigated how the causal effects of Self-Efficacy can be credibly identified using micro-economic and anthropological theory as well as state of the art statistical methods. The question whether cultural evolution might explain income differences amongst Ghana's pineapple farmers is also innovative. |
| 3 | Drought Adaptation | Why do not all farmers adapt to drought after they experienced it? | Farmers with higher self-efficacy adapt to drought, whereas others do so less or not at all | How well farmers adapt to adverse environmental conditions, such as droughts, is commonly explained with their socio-economic and institutional characteristics. When psychology and culture are investigated, the employed methods usually do not allow the kind of causal interpretation that is given by the authors. We demonstrate a more credible approach to test whether Self-Efficacy differences explain behavioral heterogeneity. |
| 4 | Persistent Constraints | Why is history differently persistent for different individuals? | Human and social capital, network effects, and exporting could all enable farmers to beat their historic prediction. | It is widely acknowledged that human and social capital are important for economic development and that history explains current income differences. Here, the two are brought together. We show that historically inherited constraints can be overcome with human and social capital. Thus, after many studies have established the commonality of historical persistence, we investigate how historical constraints can be relaxed. |
| 5 | Agricultural Training | Why is mulching widely diffused in Ghana and organic fertilizers are not, despite both being equally widely promoted? | Organic fertilizers are a more complex innovation than mulching. Thus, mulching can easily be learned from peers, whereas organic fertilizers require training. | Most studies find that farmers in developing countries benefit from trainings. Recently, it has been found that training are most effective to start the diffusion process but not to enhance it. We find that the effect of training depends critically on the nature of the trained technology. |
| 6 | UNESCO's World Heritage sites | Why are not all World Heritage sites promoting themselves as such? | It is especially economic and cultural incentives, whereas the collective brand equity and constraints are less important | This question could not be answered before, due to a prohibitively expensive data collection. With the development of a web-based big-data-collection-approach, over 300,000 respondents where surveyed. To efficiently use the available data, we use an innovative spatial econometric approach. |

The research in this thesis also raises new questions, including the following:

(1) Regarding the effect of selected cultural traits on economic outcomes, some open questions regard interaction effects between cultural traits and institutions. Especially, what kind of institutions are developed by individuals with different degrees of self-efficacy? What is the feedback from institutions on the self-efficacy of individuals? Can good institutions substitute for low self-efficacy?

(2) Above, the research on self-efficacy focuses on pineapple farmers in Ghana. Future research could explore other countries, other farmers, or other occupations such as fishers.

(3) Especially important for policy makers is the question about the malleability of self-efficacy. What is the comparative effectiveness of policy-induced self-efficacy versus naturally evolved self-efficacy? What are the most cost-effective ways to increase self-efficacy?

(4) Since we have found that site specific benefits are important to explain the marketing behavior of the World Heritage sites, a research question for the future is a systematic investigation of these benefits. Specifically, what determines the benefit of World Heritage inscription and how much of this is manageable?

# Publications and Author Contributions

Wuepper and Patry (2016). "The World Heritage list: Which sites promote the brand? A big data spatial econometrics approach." Journal of Cultural Economics: 1-21.

Marc Patry provided the data as well as ideas and feedback to the article. David Wuepper came up with the research question, conducted the analysis and wrote the article. The basic idea was to come up with a creative idea to capture how much the World Heritage sites all around the world are marketing themselves as such. Marc Patry organized a partnership with the travel-website TripAdvisor, to collect the required data decentralized and online. David Wuepper developed the research idea to explain the observed phenomenon. There was apparent heterogeneity across regions and sites, and the data allowed econometric analyzes. For the research article, much more data was needed than that collected by the authors, so David Wuepper collected additional data together with a large group of United Nations Volunteers. He also used secondary data, available from multiple sources. David Wuepper also performed all analyzed and both authors discussed the various choices to be made during the analysis, and especially what the results imply. Both authors came to the conclusion that it is more incentives than constraints that explain marketing heterogeneity. This finding is important, because it suggests that either the sites that do not market themselves as World Heritage sites must be supported to benefit or informed about their benefit, or secondly, marketing could be made mandatory. The reason why this is important, is because the World Heritage program is not only there for conservation, but also education and development. The brand value is a function of all the sites promoting themselves as World Heritage sites, and it is often exactly the sites that could contribute the most, that do not see an advantage in doing so. This research has benefited greatly from many discussion with various policy makers, when David Wuepper was an intern at the World Heritage Center of the UNESCO in Paris and Marc Patry was head of the Special Unit Forests. It also benefited tremendously from multiple rounds of peer review.

Wuepper and Sauer (2016). "Explaining the Performance of Contract Farming in Ghana: The Role of Self-Efficacy and Social Capital." Food Policy (62): 11-27.

David Wuepper came up with the research idea and the study design, performed the statistical analysis and wrote the article. Johannes Sauer improved the article with his feedback and suggestions throughout the whole process. David Wuepper developed a survey to interview a representative sample of pineapple farmers in Ghana. Several people provided invaluable help and support. These people where especially Dr. Linda Kleemann, and also Prof. Johannes Sauer, Prof. Awudu Abdulai, Barbara Drosten, several indviduals affilited with the GIZ in Ghana, and Prof. Mosche Ben-Akiva. David Wuepper then collected new survey data from 400 pineapple farmers, together with a large group of enumerators and the assistance of the GIZ. Prof. Alexander Moradi provided a rich dataset on the location, characterstics, and performance of the colonial cocoa cooperatives in southern Ghana. Several other historical datasets were publicly avilable and could be cominded with the main dataset. David Wuepper performed all econometric analyses under the guidance of Prof. Johannes Sauer and with feedback from many individuals as well as conference participants (amongst others, from the International Association of Applied Econometrics, The American Association of Agricultural Economics). Also this article benefited tremendously from peer review, and the final article is clearly better than previous versions.

Wuepper and Drosten (). "The Profitability of Investment Self-Efficacy: Agent-Based Modeling and Empirical Evidence from Rural Ghana", submitted to Journal of Evolutionary Economics.

David Wuepper collected the data, programmed the agent based model, performed the econometric analysis and wrote most of the paper. Barbara Drosten contributed ideas to the survey, and provided content, especially to the introduction and the discussion section. This research is based on the same survey data as the previous article but combined with new historical data, all publicly available. A major challenge of this research was to develop a clear concept of the main variable: Investment Self-Efficacy. David Wuepper and Barbara Drosten developed a clear concept, proposing that cultural evolution explains where it comes from (using an agent based model as rigorous, formal framework) and what it does (the agent based model produces different investment behaviors as a function of historical investment experiences). This was then tested with the survey data from Ghana's pineapple farmers. David Wuepper set the empirical model up as his agent based model had suggested. The historical experiences of the farmers were approximated by which kind of crop they optimally depended on in pre-colonial times. The final outcome of interest was income, and the proposed mechanism was investment behavior. Indeed, the data is consistent with historical experiences affecting investment choices, producing long-term income differences, between individuals and regions.

Wuepper, Zilberman, and Sauer (). “Self-Efficacy or Farming Skills: What Matters more for the Adaptive Capacity of Ghana’s Pineapple Farmers?”, submitted to Journal of Agricultural Economics.

David Wuepper and David Zilberman came up with the research question. David Wuepper also prepared and analyzed the data, and wrote the article. David Zilberman and Johannes Sauer contributed to the research with their ideas, discussion, and suggestions. During a research visit at UC Berkeley, David Wuepper and David Zilberman discussed the concept of perceived self-efficacy and how it could affect not only general behavior and outcomes, but particularly how it affects the response to adversity. Existing research could only establish an empirical relationship between reported perceptions and behavioral intentions, but until today, there is no published research, robustly identifying the causal effect of perceived self-efficacy on the response to adversity. Using the data from the Ghana survey, David Wuepper econometrically analyzed whether farmers differ in their response to drought as a function of their self-efficacy. This analysis was frequently discussed with David Zilberman and Johannes Sauer, and especially the robustness checks benefited from these discussions. This research was also presented at the CSAE conference in Oxford, and many helpful comments and discussion improved this work.

Wuepper and Sauer (). “Moving Forward in Rural Ghana: Investing in Social and Human Capital Mitigates Historical Constraints”, submitted to Economic History of Developing Regions

David Wuepper performed the analysis and wrote the article. Johannes Sauer provided valuable comments and suggestions. David Wuepper discussed with Prof. Davide Cantoni economic history research and a possible connection with agricultural development economics. Prof. Davide Cantoni suggested a general research gap. We have ample empirical evidence about historical persistence in economic development, but we know far less what allows individuals to break out of this. In this discussion David Wuepper developed the research idea to first predict the income of the sampled pineapple farmers in Ghana, using historical variables. In a second, step the idea was to analyze the outliers. Thus, David Wuepper used data on historical production systems and slavery to predict current incomes. Then, various possible explanations for relative income improvements were tested. The analyses and tests benefited from the feedback of Johannes Sauer. Also, Dr. Francesco Cinnirella and Dr. Matthias Blum gave helpful feedback and the article has received several round of peer review. It is currently in the last round of minor improvements and has especially seen improvements in the volume of robustness checks as well as literature review.

Wuepper, Sauer, and Kleemann (). “The Diffusion of Sustainable Intensification Practices in Ghana: Why is Mulching so much more Common than the use of Organic Fertilizers?”, submitted to Applied Economic Perspectives and Policy

David Wuepper collected the data, performed the analysis and wrote the article. Linda Kleemann greatly supported the data collection in Ghana and contributed her own data from 2010. Linda Kleemann and Johannes Sauer also contributed to the article with ideas, feedback, and discussions.

The research question was an idea that developed during discussion with members of the GIZ. Many actors provide agricultural trainings to the farmers in Ghana, but there is little information on whether these are effective. Anecdotal evidence was quite mixed, also in the field. Apparently some combination of farmers, trainers, and context lead to innovation, and others did not. David Wuepper performed an econometric analysis, with the help of Linda Kleemann and Johannes Sauer. Several identification and data challenges made several refinements necessary, until the results are now truly robust and reliable. The main challenge was endogeneity in regard to social networks and trainings. Using a state of the art framework, David Wuepper finds that it is mostly the characteristics of the innovation that matters. The research has benefitted from the feedback of Christopher Udry, several peer reviewers, as well discussions at the Bioecon at Cambridge, and methodological advice from Joshua Angrist.

## *Acknowledgements*

I thank my supervisor Prof. Johannes Sauer and the people who have co-authored chapters in this thesis: Barbara Drosten (chapter 2), Prof. David Zilberman (chapter 3), Linda Kleemann (chapter 5), and Marc Patry (chapter 6). I also thank everybody who supported the data collection in Ghana, including Prof. Awudu Abdulai, Linda Kleemann, the GIZ staff in Ghana, and the enumerators. Furthermore, I received great feedback on ideas from Prof. Alexander Moradi, Prof. Davide Cantoni, Prof. Martin Quaas, Dr. Francesco Cinnirella, Dr. Matthias Blum, Prof. Travis Lybbert, Prof. Mosche Ben Akiva, and Prof. Joshua Angrist. I had the opportunity to present and discuss large parts of my research on international conferences, where I received invaluable feedback (amongst others, from the international society of applied econometrics, the American association of agricultural economics, the Bioecon conference at Cambridge University). A part of this thesis has been written during two research visits at UC Berkeley in 2015 and 2016 and I thank Prof. David Zilberman for his supervision during this time. Special thanks go to my family, for their love and support.

# References

Aaker, D. A. (2009). Managing brand equity, Simon and Schuster.
Abdulai, A., V. Owusu and R. Goetz (2011). "Land tenure differences and investment in land improvement measures: Theoretical and empirical analyses." Journal of Development Economics **96**(1): 66-78.
Acemoglu, D., F. A. Gallego and J. A. Robinson (2014). "Institutions, Human Capital, and Development." Annual Review of Economics **6**(1): 875-912.
Acemoglu, D., S. Johnson and J. A. Robinson (2002). "Reversal of fortune: Geography and institutions in the making of the modern world income distribution." Quarterly Journal of economics: 1231-1294.
Acemoglu, D. and J. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation." The American Economic Review **91**(5): 1369-1401.
Acemoglu, D. and J. Robinson (2012). Why Nations Fail: The Origins of Power, Prosperity, and Poverty, Crown Business.
Adger, W. N. (2003). "3. Social Aspects of Adaptive Capacity." Climate change, adaptive capacity and development: 29.
Adger, W. N. (2010). Social capital, collective action, and adaptation to climate change. Der klimawandel, Springer**:** 327-345.
Adger, W. N., J. Barnett, K. Brown, N. Marshall and K. O'Brien (2013). "Cultural dimensions of climate change impacts and adaptation." Nature Climate Change **3**(2): 112-117.
Adger, W. N., S. Dessai, M. Goulden, M. Hulme, I. Lorenzoni, D. R. Nelson, L. O. Naess, J. Wolf and A. Wreford (2009). "Are there social limits to adaptation to climate change?" Climatic change **93**(3-4): 335-354.
Ahlerup, P., O. Olsson and D. Yanagizawa (2009). "Social capital vs institutions in the growth process." European Journal of Political Economy **25**(1): 1-14.
Akerlof, G. A. and R. Kranton (2010). "Identity economics." The Economists' Voice **7**(2).
Alesina, A. and P. Giuliano (2013). Culture and institutions, National Bureau of Economic Research.
Anderson, J. R. and G. Feder (2004). "Agricultural extension: Good intentions and hard realities." The World Bank Research Observer **19**(1): 41-60.
Angrist, J. D. (2001). "Estimation of limited dependent variable models with dummy endogenous regressors." Journal of business & economic statistics **19**(1).
Angrist, J. D. and J.-S. Pischke (2008). Mostly harmless econometrics: An empiricist's companion, Princeton university press.
Arezki, R., R. Cherif and J. Piotrowski (2009). Tourism specialization and economic development: Evidence from the UNESCO World Heritage List, International Monetary Fund.
Arslan, A., N. McCarthy, L. Lipper, S. Asfaw, A. Cattaneo and M. Kokwe (2015). "Climate Smart Agriculture? Assessing the Adaptation Implications in Zambia." Journal of Agricultural Economics.
Attanasio, O., L. Pellerano and S. P. Reyes (2009). "Building Trust? Conditional Cash Transfer Programmes and Social Capital*." Fiscal Studies **30**(2): 139-177.
Austin, G. (2005). Land, Labor and Capital in Ghana: from slave to free labor in Asante, 1807-1956, Rochester NY: University of Rochester Press.
Austin, G. (2008). "The 'reversal of fortune'thesis and the compression of history: Perspectives from African and comparative economic history." Journal of international development **20**(8): 996-1027.
Awodoyin, R., F. Ogbeide and O. Oluwole (2007). "Effects of Three Mulch Types on the Growth and Yield of Tomato (Lycopersicon esculentum Mill.) and Weed Suppression in Ibadan, Rainforest-savanna Transition Zone of Nigeria." Tropical Agricultural Research and Extension **10**: 53-60.
Bandiera, O. and I. Rasul (2006). "Social networks and technology adoption in northern mozambique*." The Economic Journal **116**(514): 869-902.

Bandura, A. (1977). "Self-efficacy: toward a unifying theory of behavioral change." Psychological review **84**(2): 191.
Bandura, A. (1995). Self-efficacy in changing societies, Cambridge university press.
Bandura, A. (1997). Self-efficacy: The exercise of control, New York: Freeman.
Bandura, A. (2001). "Social cognitive theory of mass communication." Media psychology **3**(3): 265-299.
Bandura, A. (2012). "On the functional properties of perceived self-efficacy revisited." Journal of Management **38**(1): 9-44.
Bandura, A. (2015). "On deconstructing commentaries regarding alternative theories of self-regulation." Journal of Management **41**(4): 1025-1044.
Bandura, A., C. Barbaranelli, G. V. Caprara and C. Pastorelli (2001). "Self-efficacy beliefs as shapers of children's aspirations and career trajectories." Child development: 187-206.
Banerjee, A. and E. Duflo (2011). Poor economics: A radical rethinking of the way to fight global poverty, PublicAffairs.
Banerjee, A. V. and S. Mullainathan (2008). "Limited attention and income distribution." The American Economic Review: 489-493.
Bardhan, P. and C. Udry (1999). Development microeconomics, Oxford University Press.
Barrett, C. B., M. E. Bachke, M. F. Bellemare, H. C. Michelson, S. Narayanan and T. F. Walker (2012). "Smallholder participation in contract farming: comparative evidence from five countries." World Development **40**(4): 715-730.
Barros, V., C. Field, D. Dokke, M. Mastrandrea, K. Mach, T. Bilir, M. Chatterjee, K. Ebi, Y. Estrada and R. Genova (2015). "Climate change 2014: impacts, adaptation, and vulnerability. Part B: regional aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change."
Bechara, A., H. Damasio, D. Tranel and A. R. Damasio (1997). "Deciding advantageously before knowing the advantageous strategy." Science **275**(5304): 1293-1295.
Bellemare, M. F. (2012). "As you sow, so shall you reap: The welfare impacts of contract farming." World Development **40**(7): 1418-1434.
Bellemare, M. F. (2015). "Contract Farming: What's In It for Smallholder Farmers in Developing Countries?" Choices **30**(3).
Bellemare, M. F. and L. Novak (2015). "Contract Farming and Food Security."
Bénabou, R. (2015). "The Economics of Motivated Beliefs." Revue d'économie politique **125**(5): 665-685.
Benhabib, J. and M. M. Spiegel (1994). "The role of human capital in economic development evidence from aggregate cross-country data." Journal of Monetary economics **34**(2): 143-173.
Berger, T. (2001). "Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis." Agricultural economics **25**(2-3): 245-260.
Bernard, T., S. Dercon, K. Orkin and A. S. Taffesse (2014). The future in mind: Aspirations and forward-looking behaviour in rural ethiopia. Centre for the Study of African Economies conference on economic development in Africa, Oxford, UK, March.
Bernard, T., S. Dercon, K. Orkin and A. S. Taffesse (2015). "Will Video Kill the Radio Star? Assessing the Potential of Targeted Exposure to Role Models through Video." The World Bank Economic Review **29**(suppl 1): S226-S237.
Bernard, T., S. Dercon and A. S. Taffesse (2011). "Beyond fatalism-an empirical exploration of self-efficacy and aspirations failure in Ethiopia." CSAE Working Paper 2011-03, Centre for the Study of African Economies.
Bernheim, B. D., D. Ray and S. Yeltekin (2013). Poverty and self-control, National Bureau of Economic Research.
Bertacchini, E. E. and D. Saccone (2012). "Toward a political economy of World Heritage." Journal of Cultural Economics **36**(4): 327-352.

Bertrand, M., S. Mullainathan and E. Shafir (2004). "A behavioral-economics view of poverty." American Economic Review: 419-423.
Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, J. Habarurema, A. Oduro, R. Oostendorp and C. Pattillo (2000). Exports and firm-level efficiency in African manufacturing, Centre for the Study of African Economies, University of Oxford.
Bisin, A. and T. Verdier (2001). "The economics of cultural transmission and the dynamics of preferences." Journal of Economic theory **97**(2): 298-319.
Bisin, A. and T. Verdier (2010). The economics of cultural transmission and socialization, National Bureau of Economic Research.
Blume, L. E. and S. N. Durlauf (2006). "Identifying social interactions: A review." Methods in social epidemiology **287**: 22.
Boyd, R. and P. J. Richerson (1985). "Culture and the evolutionary process." University of Chicago, Chicago.
Boyd, R., P. J. Richerson and J. Henrich (2011). "The cultural niche: Why social learning is essential for human adaptation." Proceedings of the National Academy of Sciences **108**(Supplement 2): 10918-10925.
Bramoullé, Y., H. Djebbari and B. Fortin (2009). "Identification of peer effects through social networks." Journal of econometrics **150**(1): 41-55.
Brock, W. A. and S. N. Durlauf (2001). "Discrete choice with social interactions." The Review of Economic Studies **68**(2): 235-260.
Buckley, R. (2004). "The effects of World Heritage listing on tourism to Australian national parks." Journal of Sustainable Tourism **12**(1): 70-84.
Butler, J., P. Giuliano and L. Guiso (2009). The right amount of trust, National Bureau of Economic Research.
Caeyers, B. and M. Fafchamps (2016). Exclusion Bias in the Estimation of Peer Effects. CSAE Conference, Queen's College, University of Oxford.
Cazzuffi, C. and A. Moradi (2010). Why Do Cooperatives Fail? Big versus Small in Ghanaian Cocoa Producers’ Societies, 1930-36, Centre for the Study of African Economies, University of Oxford.
Cazzuffi, C. and A. Moradi (2012). "Membership Size and Cooperative Performance: Evidence from Ghanaian Cocoa Producers' Societies, 1930–36." Economic History of Developing Regions **27**(1): 67-92.
Chanda, A. and L. Putterman (2007). "Early Starts, Reversals and Catch-up in the Process of Economic Development*." The Scandinavian Journal of Economics **109**(2): 387-413.
Chaney, N. W., J. Sheffield, G. Villarini and E. F. Wood (2014). "Development of a high-resolution gridded daily meteorological dataset over sub-saharan Africa: Spatial analysis of trends in climate extremes." Journal of Climate **27**(15): 5815-5835.
Chiapa, C., J. L. Garrido and S. Prina (2012). "The effect of social programs and exposure to professionals on the educational aspirations of the poor." Economics of Education Review **31**(5): 778-798.
Claridge, W. W. (1915). A History of the Gold Coast and Ashanti... With an introduction by Sir Hugh Clifford... With maps, John Murray.
Cogneau, D. and A. Moradi (2011). "Borders that divide: education and religion in Ghana and Togo since colonial times." Department of Economics, University of Sussex Working Paper **2911**.
Cogneau, D. and A. Moradi (2014). "Borders that divide: Education and religion in Ghana and Togo since colonial times." The Journal of Economic History **74**(03): 694-729.
Conley, T. G. and C. R. Udry (2010). "Learning about a new technology: Pineapple in Ghana." The American Economic Review: 35-69.
Cornes, R. and T. Sandler (1984). "Easy riders, joint production, and public goods." The Economic Journal: 580-598.
Cosmides, L. and J. Tooby (1994). "Better than rational: Evolutionary psychology and the invisible hand." The American Economic Review: 327-332.
Dalton, P. S., S. Ghosal and A. Mani (2015). "Poverty and aspirations failure." The Economic Journal.

Dalton, P. S., S. Ghosal and A. Mani (2016). "Poverty and Aspirations Failure." The Economic Journal **126**(590): 165-188.
Damasio, H., T. Grabowski, R. Frank, A. M. Galaburda and A. R. Damasio (1994). "The return of Phineas Gage: clues about the brain from the skull of a famous patient." Science **264**(5162): 1102-1105.
Datta, S. and S. Mullainathan (2014). "Behavioral design: A new approach to development policy." Review of Income and Wealth **60**(1): 7-35.
David, P. A. (1975). Technical choice innovation and economic growth: essays on American and British experience in the nineteenth century, Cambridge University Press.
Delavande, A., X. Giné and D. McKenzie (2011). "Measuring subjective expectations in developing countries: A critical review and new evidence." Journal of Development Economics **94**(2): 151-163.
Dell, M., B. F. Jones and B. A. Olken (2012). "Temperature shocks and economic growth: Evidence from the last half century." American Economic Journal: Macroeconomics: 66-95.
Dercon, S. and L. Christiaensen (2011). "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia." Journal of development economics **96**(2): 159-173.
Dercon, S., D. O. Gilligan, J. Hoddinott and T. Woldehanna (2009). "The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages." American Journal of Agricultural Economics **91**(4): 1007-1021.
Dercon, S. and D. Gollin (2014). "Agriculture in African development: theories and strategies." Annu. Rev. Resour. Econ. **6**(1): 471-492.
Dercon, S. and P. Krishnan (2009). "Poverty and the psychosocial competencies of children: evidence from the Young Lives sample in four developing countries." Children Youth and Environments **19**(2): 138-163.
Dercon, S. and A. Sánchez (2013). "Height in mid childhood and psychosocial competencies in late childhood: Evidence from four developing countries." Economics & Human Biology **11**(4): 426-432.
Dercon, S. and A. Singh (2013). "From nutrition to aspirations and self-efficacy: gender bias over time among children in four countries." World Development **45**: 31-50.
Di Falco, S. (2014). "Adaptation to climate change in Sub-Saharan agriculture: assessing the evidence and rethinking the drivers." European Review of Agricultural Economics **41**(3): 405-430.
Di Falco, S. and E. Bulte (2011). "A dark side of social capital? Kinship, consumption, and savings." Journal of Development Studies **47**(8): 1128-1151.
Di Falco, S., M. Veronesi and M. Yesuf (2011). "Does adaptation to climate change provide food security? A micro-perspective from Ethiopia." American Journal of Agricultural Economics **93**(3): 829-846.
Diao, X. and D. B. Sarpong (2007). Cost implications of agricultural land degradation in Ghana: an economywide, multimarket model assessment, International Food Policy Research Institute, Ghana Strategy Support Program.
Dinar, A., R. Hassan, R. Mendelsohn and J. Benhin (2012). Climate change and agriculture in Africa: impact assessment and adaptation strategies, Routledge.
Dong, Y. and A. Lewbel (2015). "A simple estimator for binary choice models with endogenous regressors." Econometric Reviews **34**(1-2): 82-105.
Duflo, E., M. Kremer and J. Robinson (2009). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya, National Bureau of Economic Research.
Duflo, E., M. Kremer and J. Robinson (2011). "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." The American Economic Review **101**(6): 2350-2390.
Dzomeku, I., G. Mahunu, T. Bayorbor and P. Obeng-Danso (2009). "Effects of mulching on weed control and yield of hot pepper and tomato in the Guinea Savannah zone." Ghana Journal of Horticulture **7**: 53-62.
Eaton, C. and A. Shepherd (2001). Contract farming: partnerships for growth, Food & Agriculture Org.

Erenstein, O. (2003). "Smallholder conservation farming in the tropics and sub-tropics: a guide to the development and dissemination of mulching with crop residues and cover crops." Agriculture, Ecosystems & Environment **100**(1): 17-37.
Fafchamps, M. (1996). "The enforcement of commercial contracts in Ghana." World Development **24**(3): 427-448.
Farmer, J. D. and D. Foley (2009). "The economy needs agent-based modelling." Nature **460**(7256): 685-686.
Feder, G., R. E. Just and D. Zilberman (1985). "Adoption of agricultural innovations in developing countries: A survey." Economic development and cultural change: 255-298.
Feder, G. and D. L. Umali (1993). "The adoption of agricultural innovations: a review." Technological forecasting and social change **43**(3): 215-239.
Feigenberg, B., E. Field and R. Pande (2013). "The economic returns to social interaction: Experimental evidence from microfinance." The Review of Economic Studies **80**(4): 1459-1483.
Feigenberg, B., E. M. Field and R. Pande (2010). Building social capital through microfinance, National Bureau of Economic Research.
Fenske, J. (2011). "The causal history of Africa: Replies to Jerven and Hopkins." Economic History of Developing Regions **26**(2): 125-131.
Fenske, J. (2011). "Land tenure and investment incentives: Evidence from West Africa." Journal of Development Economics **95**(2): 137-156.
Fernandez, A., M. Della Giusta and U. S. Kambhampati (2015). "The Intrinsic Value of Agency: The Case of Indonesia." World Development **70**: 92-107.
Flechtner, S. (2014). "Aspiration traps: When poverty stifles hope." Inequality in Focus **2**(4): 1-4.
Floyd, D. L., S. Prentice-Dunn and R. W. Rogers (2000). "A meta-analysis of research on protection motivation theory." Journal of applied social psychology **30**(2): 407-429.
Fold, N. and K. V. Gough (2008). "From smallholders to transnationals: the impact of changing consumer preferences in the EU on Ghana’s pineapple sector." Geoforum **39**(5): 1687-1697.
Foster, A. D. and M. R. Rosenzweig (1995). "Learning by doing and learning from others: Human capital and technical change in agriculture." Journal of political Economy: 1176-1209.
Foster, A. D. and M. R. Rosenzweig (2010). "Microeconomics of technology adoption." Annual Review of Economics **2**.
Frey, B. S., P. Pamini and L. Steiner (2011). "What determines the world heritage list? An econometric analysis." An Econometric Analysis (January 1, 2011). University of Zurich Department of Economics Working Paper(1).
Fukuyama, F. (2001). "Social capital, civil society and development." Third world quarterly **22**(1): 7-20.
Gallup, J. L., J. D. Sachs and A. D. Mellinger (1999). "Geography and economic development." International regional science review **22**(2): 179-232.
Galor, O. and S. Michalopoulos (2012). "Evolution and the growth process: Natural selection of entrepreneurial traits." Journal of Economic Theory **147**(2): 759-780.
Galor, O. and Ö. Özak (2014). The Agricultural Origins of Time Preference, National Bureau of Economic Research.
Gatune, J., M. Chapman-Kodam, K. Korboe, F. Mulangu and M. Raktoarisoa (2013). Analysis of Trade Impacts on the Fresh Pineapple Sector in Ghana. FAO Commodity and Trade Policy Research Working Paper No.41. Rome, Food and Agricultural Organization of the United Nations.
Gebrehiwot, T. and A. van der Veen (2015). "Farmers prone to drought risk: why some farmers undertake farm-level risk-reduction measures while others not?" Environmental management **55**(3): 588-602.
Gennaioli, N., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2013). "Human Capital and Regional Development*." Quarterly Journal of Economics **128**(1).
German Society for International Cooperation (2005). Market Oriented Agriclture Programme (MOAP).

Giavazzi, F., I. Petkov and F. Schiantarelli (2014). Culture: Persistence and evolution, National Bureau of Economic Research.
Gigerenzer, G. and W. Gaissmaier (2011). "Heuristic decision making." Annual review of psychology **62**: 451-482.
Gilboa, I., A. Postlewaite and D. Schmeidler (2012). "Rationality of belief." Synthese **187**(1): 11-31.
Goldstein, M. and C. Udry (2008). "The profits of power: Land rights and agricultural investment in Ghana." Journal of Political Economy **116**(6): 981-1022.
Government of Ghana (2010). Medium-term national development policy framework: Ghana shared growth and development agenda, 2010 - 2013. **1**.
Grosh, B. (1994). "Contract Farming in Africa: An Application of the New Institutional Economics." Journal of African Economies **3**(2): 231-261.
Grosjean, P. (2014). "A history of violence: The culture of honor and homicide in the US South." Journal of the European Economic Association **12**(5): 1285-1316.
Grothmann, T. and A. Patt (2005). "Adaptive capacity and human cognition: the process of individual adaptation to climate change." Global Environmental Change **15**(3): 199-213.
Guiso, L., P. Sapienza and L. Zingales (2006). "Does Culture Affect Economic Outcomes?" The Journal of Economic Perspectives **20**(2): 23-48.
Guiso, L., P. Sapienza and L. Zingales (2008). "Alfred marshall lecture social capital as good culture." Journal of the European Economic Association **6**(2-3): 295-320.
Guiso, L., P. Sapienza and L. Zingales (2010). Civic capital as the missing link, National Bureau of Economic Research.
Guryan, J., K. Kroft and M. J. Notowidigdo (2009). "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." American Economic Journal: Applied Economics **1**(4): 34-68.
Hanna, R., S. Mullainathan and J. Schwartzstein (2014). "Learning through noticing: Theory and evidence from a field experiment." The Quarterly Journal of Economics **129**(3): 1311-1353.
Harari, Y. N. (2014). Sapiens: A brief history of Humankind, Random House.
Hardiman, N. and S. Burgin (2013). "World Heritage Area listing of the Greater Blue Mountains—Did it make a difference to visitation?" Tourism Management Perspectives **6**: 63-64.
Harou, A. and T. Walker (2012). "The survival of smallholder farmers in agricultural export markets." Development **37**(11): 1717-1727.
Harrison, L. E. and S. P. Huntington (2000). Culture matters: How values shape human progress, Basic Books.
Haushofer, J. and E. Fehr (2014). "On the psychology of poverty." Science **344**(6186): 862-867.
Heisey, P. W. and W. Mwangi (1996). "Fertilizer use and maize production in sub-Saharan Africa."
Henrich, J., R. Boyd and P. J. Richerson (2008). "Five misunderstandings about cultural evolution." Human Nature **19**(2): 119-137.
Henrich, J., R. Boyd, P. Young, K. McCabe, W. Alberts, A. Ockenfelds and G. Gigerenzer (2001). "What is the role of culture in bounded rationality." Bounded rationality: the adaptive toolbox: 343-359.
Hertel, T. W. and D. B. Lobell (2014). "Agricultural adaptation to climate change in rich and poor countries: current modeling practice and potential for empirical contributions." Energy Economics **46**: 562-575.
Huang, C.-H., J.-R. Tsaur and C.-H. Yang (2012). "Does world heritage list really induce more tourists? Evidence from Macau." Tourism Management **33**(6): 1450-1457.
Huillery, E. (2009). "History matters: The long-term impact of colonial public investments in French West Africa." American Economic Journal: Applied Economics: 176-215.
Huillery, E. (2011). "The Impact of European Settlement within French West Africa: Did pre-colonial prosperous areas fall behind?" Journal of African Economies **20**(2): 263-311.

Hutter, M. and D. Throsby (2008). Beyond price: Value in culture, economics, and the arts, Cambridge University Press.
Inglehart, R. and W. E. Baker (2000). "Modernization, cultural change, and the persistence of traditional values." American sociological review: 19-51.
Janssen, M. A. (2005). "Agent-based modelling." Modelling in ecological economics: 155-172.
Jedwab, R. and A. Moradi (2012). Colonial investments and long term development in Africa: evidence from Ghanaian railroads. a number of conferences and seminars including Oxford, Bocconi, LSE and George Washington University.
Jedwab, R. and A. Moradi (2015). "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." Review of Economics and Statistics.
Jensen, R. and E. Oster (2009). "The power of TV: Cable television and women's status in India." The Quarterly Journal of Economics **124**(3): 1057-1094.
Jimura, T. (2011). "The impact of world heritage site designation on local communities–A case study of Ogimachi, Shirakawa-mura, Japan." Tourism Management **32**(2): 288-296.
Jones, C. and M. Munday (2001). "Blaenavon and United Nations World Heritage Site status: is conservation of industrial heritage a road to local economic development?" Regional Studies **35**(6): 585-590.
Jones, L. and E. Boyd (2011). "Exploring social barriers to adaptation: insights from Western Nepal." Global Environmental Change **21**(4): 1262-1274.
Just, D. R. (2002). Information, processing capacity, and judgment bias in risk assessment. A Comprehensive Assessment of the Role of Risk in US Agriculture, Springer: 81-101.
Just, D. R., Y. Cao and D. Zilberman (2009). Risk, Overconfidence and Production in a Competitive Equilibrium. 2009 Annual Meeting, July 26-28, 2009, Milwaukee, Wisconsin, Agricultural and Applied Economics Association.
Kabunga, N. S., T. Dubois and M. Qaim (2012). "Heterogeneous information exposure and technology adoption: The case of tissue culture bananas in Kenya." Agricultural Economics **43**(5): 473-486.
Karlan, D., R. D. Osei, I. Osei-Akoto and C. Udry (2012). Agricultural decisions after relaxing credit and risk constraints, National Bureau of Economic Research.
Karlan, D. S. (2005). "Using experimental economics to measure social capital and predict financial decisions." American Economic Review: 1688-1699.
Kaul, I., I. Grunberg and M. Stern (1999). Global public goods: international cooperation in the 21st century, Oxford University Press.
Kautonen, T., M. Van Gelderen and E. T. Tornikoski (2013). "Predicting entrepreneurial behaviour: a test of the theory of planned behaviour." Applied Economics **45**(6): 697-707.
Kayahan, B. and B. Vanblarcom (2012). "Cost Benefit Analysis of UNESCO World Heritage Site Designation in Nova Scotia." Review of Economic Analysis **4**(2): 247-273.
King, L. M. (2011). Investigating the role of the World Heritage brand in attracting visitors to protected areas in Queensland, Australia, James Cook University.
King, L. M. and E. A. Halpenny (2014). "Communicating the World Heritage brand: visitor awareness of UNESCO's World Heritage symbol and the implications for sites, stakeholders and sustainable management." Journal of Sustainable Tourism(ahead-of-print): 1-19.
Kirsten, J. and K. Sartorius (2002). "Linking agribusiness and small-scale farmers in developing countries: is there a new role for contract farming?" Development Southern Africa **19**(4): 503-529.
Kleemann, L. and A. Abdulai (2012). Organic certification, agro-ecological practices and return on investment: Farm level evidence from Ghana, Kiel Working Paper.
Kleemann, L. and A. Abdulai (2013). "Organic certification, agro-ecological practices and return on investment: Evidence from pineapple producers in Ghana." Ecological Economics **93**: 330-341.
Kleemann, L., A. Abdulai and M. Buss (2014). "Certification and access to export markets: Adoption and return on investment of organic-certified pineapple farming in Ghana." World Development **64**: 79-92.

Knack, S. and P. Keefer (1997). "Does social capital have an economic payoff? A cross-country investigation." The Quarterly journal of economics: 1251-1288.
Knack, S. and P. J. Zak (2003). "Building trust: public policy, interpersonal trust, and economic development." Sup. Ct. Econ. Rev. **10**: 91.
Krishnan, P. and S. Krutikova (2013). "Non-cognitive skill formation in poor neighbourhoods of urban India." Labour Economics **24**: 68-85.
Krishnan, P. and M. Patnam (2014). "Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption?" American Journal of Agricultural Economics **96**(1): 308-327.
Kroemker, D. and H.-J. Mosler (2002). "Human vulnerability—factors influencing the implementation of prevention and protection measures: an agent based approach." Global Environmental Change in Alpine Regions. Impact, Recognition, Adaptation, and Mitigation. Edward Elgar, Cheltenham: 95-114.
Krueger, N. and P. R. Dickson (1994). "How Believing in Ourselves Increases Risk Taking: Perceived Self-Efficacy and Opportunity Recognition." Decision Sciences **25**(3): 385-400.
La Ferrara, E., A. Chong and S. Duryea (2012). "Soap operas and fertility: Evidence from Brazil." American Economic Journal: Applied Economics: 1-31.
La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. W. Vishny (1997). "Trust in Large Organizations." The American Economic Review: 333-338.
Laajaj, R. (2014). "Closing the Eyes on a Gloomy Future: Psychological Causes and Economic Consequences."
LeSage, J. and R. K. Pace (2010). Introduction to spatial econometrics, CRC press.
LeSage, J. P. (1999). Spatial econometrics, Regional Research Institute, West Virginia University.
Lewbel, A., Y. Dong and T. T. Yang (2012). "Comparing features of convenient estimators for binary choice models with endogenous regressors." Canadian Journal of Economics/Revue canadienne d'économique **45**(3): 809-829.
Licciardi, G. and R. Amirtahmasebi (2012). The Economics of Uniqueness: Investing in Historic City Cores and Cultural Heritage Assets for Sustainable Development, World Bank Publications.
Lovejoy, P. E. (2011). Transformations in slavery: a history of slavery in Africa, Cambridge University Press.
Macdonald, G. (1898). The Gold Coast, Past and Present.
Maddux, J. E. (1995). Self-efficacy theory, Springer.
Maertens, A. and C. B. Barrett (2013). "Measuring Social Networks' Effects on Agricultural Technology Adoption." American Journal of Agricultural Economics **95**(2): 353-359.
Magnan, N., D. J. Spielman, T. J. Lybbert and K. Gulati (2015). "Leveling with Friends: Social Networks and Indian Farmers' Demand for a Technology with Heterogeneous Benefits." Journal of Development Economics.
Mani, A., S. Mullainathan, E. Shafir and J. Zhao (2013). "Poverty impedes cognitive function." science **341**(6149): 976-980.
Mankiw, G., D. Romer and D. N. Weil (1992). "A contribution to the empirics of economic growth." Quarterly Journal of Economics **107**(2): 407-437.
Manski, C. F. (1993). "Identification of endogenous social effects: The reflection problem." The review of economic studies **60**(3): 531-542.
Manski, C. F. (2000). Economic analysis of social interactions, National bureau of economic research.
Marcotte, P. and L. Bourdeau (2012). "Is the World Heritage label used as a promotional argument for sustainable tourism?" Journal of Cultural Heritage Management and Sustainable Development **2**(1): 80-91.
Mas, F. and J. Nicolau (2010). Contribution of individual to collective brands, Instituto Valenciano de Investigaciones Económicas, SA (Ivie).
Mazza, I. (2011). Public choice. Handbook on Cultural Economics. R. Towse. Cheltenam, Edward Elgar**:** 362 - 389.

McIntyre, B., H. Herren, J. Wakhungu and R. Watson (2009). "Agriculture at a Crossroads: Sub-Saharan Africa." Science And Technology **5**.
McMillan, M. (2012). "Blue Skies: How One Firm Overcame “Binding” Constraints."
Meijerink, G., E. Bulte and D. Alemu (2014). "Formal institutions and social capital in value chains: The case of the Ethiopian Commodity Exchange." Food Policy **49**: 1-12.
Mesoudi, A. (2016). "Cultural evolution: integrating psychology, evolution and culture." Current Opinion in Psychology **7**: 17-22.
Michalopoulos, S. and E. Papaioannou (2012). "National institutions and subnational development in Africa." The quarterly journal of economics **129**(1): 151-213.
Michalopoulos, S., L. Putterman and D. N. Weil (2016). "The Influence of Ancestral Lifeways on Individual Economic Outcomes in Sub-Saharan Africa." NBER Working Paper 21907.
Millenium Development Authority (2011). Millenium Challenge Account Ghana Program Agriculture Project.
Moffitt, R. A. (2001). "Policy interventions, low-level equilibria, and social interactions." Social dynamics **4**(45-82): 6-17.
Morgan, T. J., C. P. Cross and L. E. Rendell (2015). Nothing in Human Behavior Makes Sense Except in the Light of Culture: Shared Interests of Social Psychology and Cultural Evolution. Evolutionary Perspectives on Social Psychology, Springer: 215-228.
Moser, C. M. and C. B. Barrett (2006). "The complex dynamics of smallholder technology adoption: the case of SRI in Madagascar." Agricultural Economics **35**(3): 373-388.
Moya, A. and M. Carter (2014). Violence and the Formation of Hopelessness and Pessimistic Prospects of Upward Mobility in Colombia, National Bureau of Economic Research.
Mullainathan, S. (2005). Development economics through the lens of psychology. Annual World Bank Conference in Development Economics 2005: Lessons of Experience.
Mullainathan, S. and E. Shafir (2013). Scarcity: Why having too little means so much, Macmillan.
Munshi, K. (2004). "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." Journal of Development Economics **73**(1): 185-213.
Murdock, G. P. (1959). "Africa: its peoples and their culture history."
Murdock, G. P. (1967). "Ethnographic atlas."
Nicholson, N., E. Soane, M. Fenton-O'Creevy and P. Willman (2005). "Personality and domain-specific risk taking." Journal of Risk Research **8**(2): 157-176.
Nicolau, J. and F. Mas (2014). "Detecting Free Riders in Collective Brands through a Hierarchical Choice Process." Journal of Travel Research: 0047287513517419.
Norman, J. (1986). "Effects of mulching and nitrogen fertilization on" Sugarloaf" Pineapple, Ananas comosus (L) Merr." Der Tropenlandwirt-Journal of Agriculture in the Tropics and Subtropics **87**(1): 47-53.
Norris, P. (2009). Democracy Cross-national Data, Release 3.0 Spring 2009. H. K. School.
Nunn, N. (2007). The long-term effects of Africa's slave trades, National Bureau of Economic Research.
Nunn, N. (2008). "The Long-Term Effects of Africa's Slave Trades." The Quarterly Journal of Economics **123**(1): 139-176.
Nunn, N. (2009). The importance of history for economic development, National Bureau of Economic Research.
Nunn, N. (2010). "Religious conversion in colonial Africa." The American Economic Review: 147-152.
Nunn, N. (2012). "Culture and the historical process." Economic History of Developing Regions **27**(sup1): S108-S126.
Nunn, N. (2013). "Historical development." Handbook of Economic Growth.
Nunn, N. and D. Puga (2012). "Ruggedness: The blessing of bad geography in Africa." Review of Economics and Statistics **94**(1): 20-36.
Nunn, N. and L. Wantchekon (2011). "The Slave Trade and the Origins of Mistrust in Africa." The American Economic Review **101**: 3221-3252.

Obikili, N. (2015). "Social Capital and Human Capital in the Colonies: A Study of Cocoa Farmers in Western Nigeria." Economic History of Developing Regions **30**(1): 1-22.
Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action, Cambridge university press.
Oya, C. (2012). "Contract Farming in Sub-Saharan Africa: A Survey of Approaches, Debates and Issues." Journal of Agrarian Change **12**(1): 1-33.
Pajares, F. (1997). "Current directions in self-efficacy research." Advances in motivation and achievement **10**(149): 1-49.
Pajares, F. (2002). Overview of social cognitive theory and of self-efficacy.
Pamuk, H., E. Bulte, A. Adekunle and A. Diagne (2014). "Decentralised innovation systems and poverty reduction: experimental evidence from Central Africa." European Review of Agricultural Economics: jbu007.
Pamuk, H., E. Bulte and A. A. Adekunle (2014). "Do decentralized innovation systems promote agricultural technology adoption? Experimental evidence from Africa." Food Policy **44**: 227-236.
Pan, Y., S. C. Smith and M. Sulaiman (2015). Agricultural Extension and Technology Adoption for Food Security: Evidence from Uganda, IZA Discussion Papers.
Papagapitos, A. and R. Riley (2009). "Social trust and human capital formation." Economics Letters **102**(3): 158-160.
Pasquier-Doumer, L. and F. R. Brandon (2015). "Aspiration failure: a poverty trap for indigenous children in Peru?" World Development **72**: 208-223.
Petrin, A. and K. Train (2010). "A control function approach to endogeneity in consumer choice models." Journal of Marketing Research **47**(1): 3-13.
Pindyck, R. S. (2013). "Climate Change Policy: What Do the Models Tell Us?" Journal of Economic Literature **51**(3): 860-872.
Pisati, M. (2001). "sg162: tools for spatial data analysis." Stata Technical Bulletin **60**: 21-37.
Pretty, J., C. Toulmin and S. Williams (2011). "Sustainable intensification in African agriculture." International journal of agricultural sustainability **9**(1): 5-24.
Pritchett, L. and S. Kapoor (2009). Moving out of poverty: Success from the bottom up, World Bank Publications.
Putnam, R. D., R. Leonardi and R. Y. Nanetti (1994). Making democracy work: Civic traditions in modern Italy, Princeton university press.
Rao, E. J. and M. Qaim (2011). "Supermarkets, farm household income, and poverty: insights from Kenya." World Development **39**(5): 784-796.
Rebanks, J. (2009). World Heritage Status: Is there opportunity for economic gain? Research and analysis of the socio-economic impact potential of UNESCO World Heritage Site status, Rebanks Consulting Ltd and Trends Business Research Ltd on behalf of the Lake District World Heritage Project.
Rees, D., A. Westby, K. Tomlins, Q. Van Oirschot, M. U. Cheema, E. Cornelius and M. Amjad (2012). "Tropical root crops." Crop Post-Harvest: Science and Technology, Perishables **3**: 392.
Reuben, E., P. Sapienza and L. Zingales (2009). "Is mistrust self-fulfilling?" Economics Letters **104**(2): 89-91.
Richerson, P. J. and R. Boyd (2008). Not by genes alone: How culture transformed human evolution, University of Chicago Press.
Rogers, E. M. (2010). Diffusion of innovations, Simon and Schuster.
Rosenzweig, M. R. and K. I. Wolpin (2000). "Natural" natural experiments" in economics." Journal of Economic Literature **38**(4): 827-874.
Ryan, J. and S. Silvanto (2011). "A brand for all the nations: The development of the World Heritage Brand in emerging markets." Marketing Intelligence & Planning **29**(3): 305-318.
Sachs, J. D. (2003). Institutions don't rule: direct effects of geography on per capita income, National Bureau of Economic Research.

Sauer, J. and D. Zilberman (2012). "Sequential technology implementation, network externalities, and risk: the case of automatic milking systems." Agricultural Economics **43**(3): 233-252.
Schultz, T. W. (1975). "The value of the ability to deal with disequilibria." Journal of economic literature: 827-846.
Schultz, T. W. (1980). "Nobel lecture: The economics of being poor." The Journal of Political Economy: 639-651.
Schwarzer, R. (2014). Self-efficacy: Thought control of action, Taylor & Francis.
Shah, A. K., S. Mullainathan and E. Shafir (2012). "Some consequences of having too little." Science **338**(6107): 682-685.
Simon, H. A. (1982). Models of bounded rationality: Empirically grounded economic reason, MIT press.
Smit, B. and O. Pilifosova (2003). "Adaptation to climate change in the context of sustainable development and equity." Sustainable Development **8**(9): 9.
Smit, B. and J. Wandel (2006). "Adaptation, adaptive capacity and vulnerability." Global environmental change **16**(3): 282-292.
Smith, M. (2002). "A critical evaluation of the global accolade: the significance of World Heritage Site status for Maritime Greenwich." International Journal of Heritage Studies **8**(2): 137-151.
Snapp, S. and B. Pound (2011). Agricultural Systems: Agroecology and Rural Innovation for Development: Agroecology and Rural Innovation for Development, Academic Press.
Spolaore, E. and R. Wacziarg (2013). "How Deep Are the Roots of Economic Development?" Journal of Economic Literature **51**(2): 325-369.
Suri, T. (2011). "Selection and comparative advantage in technology adoption." Econometrica **79**(1): 159-209.
Suzuki, A., L. S. Jarvis and R. J. Sexton (2011). "Partial vertical integration, risk shifting, and product rejection in the high-value export supply chain: The Ghana pineapple sector." World Development **39**(9): 1611-1623.
Tabellini, G. (2008). "The Scope of Cooperation: Values and Incentives." The Quarterly Journal of Economics **123**(3): 905-950.
Tabellini, G. (2010). "Culture and institutions: economic development in the regions of Europe." Journal of the European Economic Association **8**(4): 677-716.
Tafere, Y. (2014). "Education aspirations and barriers to achievement for young people in Ethiopia." Young Lives Working Paper(120).
Talhelm, T., X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan and S. Kitayama (2014). "Large-scale psychological differences within China explained by rice versus wheat agriculture." Science **344**(6184): 603-608.
Teorell, J. D., Stefan ; Holmberg, Sören; Rothstein, Bo ; Hartmann, Felix ; Svensson, Richard (2015). The Quality of Government Standard Dataset, version Jan15, University of Gothenburg: The Quality of Government Institute, .
Throsby, D., G. Licciardi and R. Amirtahmasebi (2012). "Heritage economics: a conceptual framework." The Economics of Uniqueness: 45.
Throsby, D. and A. Zednik (2013). "The Economic and cultural value of paintings: some empirical evidence." Handbook of the Economics of Art and Culture **2**: 81.
Tilman, D., C. Balzer, J. Hill and B. L. Befort (2011). "Global food demand and the sustainable intensification of agriculture." Proceedings of the National Academy of Sciences **108**(50): 20260-20264.
Tisdell, C. and C. Wilson (2002). "World heritage listing of Australian natural sites: tourism stimulus and its economic value." Economic analysis and policy **32**: 27-50.
Tol, R. S., S. Fankhauser and J. B. Smith (1998). "The scope for adaptation to climate change: what can we learn from the impact literature?" Global Environmental Change **8**(2): 109-123.
Udry, C. and S. Anagol (2006). "The return to capital in Ghana." The American Economic Review: 388-393.

Udry, C. R. and T. G. Conley (2004). "Social Networks in Ghana."
UNESCO (1972). Convention Concerning the Protection of the World Cultural and Natural Heritage. N. Adopted by the General Conference at its Seventeenth Session Paris. Paris, United Nations Educational, Scientific and Cultural Organisation, Paris, France.
UNESCO (2013). Operational Guidelines for the Implementation of the World Heritage Convention. S. A. C. O. UNITED NATIONS EDUCATIONAL. Paris.
USAID (2007). Trade and Investment Program for a Competitive Export Economy (TIPCEE) - Forth Year Work Plan for Partners. October 2007 - September 2008.
USAID (2009). Trade and Investment Program for a Competitive Export Economy - Final Report.
USAID (2013). Agricultural Development and Value Chain Enhancement (ADVANCE) - Project profile.
Van Biesebroeck, J. (2005). "Exporting raises productivity in sub-Saharan African manufacturing firms." Journal of International economics **67**(2): 373-391.
Van den Broeck, K. and S. Dercon (2011). "Information flows and social externalities in a Tanzanian banana growing village." The journal of development studies **47**(2): 231-252.
van Duinen, R., T. Filatova, P. Geurts and A. van der Veen (2014). "Coping with drought risk: empirical analysis of farmers' drought adaptation in the south-west Netherlands." Regional Environmental Change **15**(6): 1081-1093.
Voigtländer, N. and H.-J. Voth (2012). "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany*." The Quarterly Journal of Economics **127**(3): 1339-1392.
Wagner, J. (2007). "Exports and productivity: A survey of the evidence from firm-level data." The World Economy **30**(1): 60-82.
Wantchekon, L., N. Novta and M. Klašnja (2015). "Education and Human Capital Externalities: Evidence from Colonial Benin." The Quarterly Journal of Economics **January 30, 2015**.
Ward, M. D. and K. S. Gleditsch (2008). Spatial regression models, Sage.
Ward, W. E. F. (1966). A history of Ghana, Allen & Unwin.
Weber, E. U., A.-R. Blais and N. E. Betz (2002). "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors." Journal of behavioral decision making **15**: 263-290.
Weinberg, R. S., D. Gould, D. Yukelson and A. Jackson (1981). "The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task." Journal of Sport Psychology.
Will, M. (2013). Contract Farming Handbook - A practical guide for linking small-scale producers and buyers trhough business model innovation, Deutsche Gesellschaft für International Zusammenarbeit.
Woodberry, R. D. (2004). The shadow of empire: Christian missions, colonial policy, and democracy in postcolonial societies, University of North Carolina at Chapel Hill.
Woolcock, M. and D. Narayan (2000). "Social capital: Implications for development theory, research, and policy." The world bank research observer **15**(2): 225-249.
World Bank (2011). Horticultural Exports from Ghana: A Strategic Study. Discussion Paper.
Wuepper, D. (2014). Opportunities and constraints to integrate Ghana's smallholder farmers into global values chains. African Economic Outlook 2014, African Development Bank, Development Centre of the Organisation for Economic Cooperation and Development (OECD), United Nations Development Programme (UNDP).
Wuepper, D. and B. Drosten (2016). Historical Return on Investment and Current Economic Outcomes: The Cultural Evolution of Investment Self-Efficacy Working Paper August 2016.
Wuepper, D. and M. Patry (2016). "The World Heritage list: Which sites promote the brand? A big data spatial econometrics approach." Journal of Cultural Economics: 1-21.
Wuepper, D. and J. Sauer (2016). "Explaining the Performance of Contract Farming in Ghana: The Role of Self-Efficacy and Social Capital." Food Policy(62): 11-27.
Wuepper, D. and J. Sauer (2016). Moving Forward in Rural Ghana: Investing in Social and Human Capital Mitigates Historical Constraints, forthcoming.

Wuepper, D., J. Sauer and L. Kleemann (2014). "Sustainable Intensification of Pineapple Farming in Ghana: Training and Complexity." Working Paper Kiel Institute for the World Economy.
Wuepper, D., D. Zilberman and J. Sauer (2016). Self-Efficacy or Farming Skills: What matters more for the Adaptive Capacity of Ghana's Pineapple Farmers? Working Paper April 2016.
Yang, C.-H., H.-L. Lin and C.-C. Han (2010). "Analysis of international tourist arrivals in China: The role of World Heritage Sites." Tourism management **31**(6): 827-837.
Yohe, G. and R. S. Tol (2002). "Indicators for social and economic coping capacity—moving toward a working definition of adaptive capacity." Global Environmental Change **12**(1): 25-40.
Zak, P. J. and S. Knack (2001). "Trust and growth." The economic journal **111**(470): 295-321.
Zeitlin, A. (2011). Identification and estimation of peer effects on endogenous affiliation networks: an application to Ghanaian agriculture, Working paper.
Zilberman, D., J. Zhao and A. Heiman (2012). "Adoption versus adaptation, with emphasis on climate change." Annu. Rev. Resour. Econ. **4**(1): 27-53.

# Curriculum Vitae David Wuepper

*General Information*

| | |
|---|---|
| Date of birth | 12.11.1986 |
| Place of birth | Bonn |
| Address | Technical University Munich, WZW, Alte Akademie 14, D-85354 Freising, Germany |
| Email | david.wuepper@tum.de |
| Website | http://pur.wzw.tum.de/David-Wuepper |

*Education*

| | |
|---|---|
| 2013-2016 | PhD candidate, lecturing and research assistant in Agricultural Economics - Technical University Munich – Professor Johannes Sauer |
| 2010 - 2012 | Master of Science in Agricultural Economics<br>Christian Albrecht University of Kiel |
| 2007 – 2010 | Bachelor of Science in Agricultural and Environmental Science<br>Christian Albrecht University of Kiel |

*Peer Reviewed Publications*

- "Explaining the Performance of Contract Farming in Ghana: The Role of Self-Efficacy and Social Capital." Food Policy (62): 11-27. With Johannes Sauer
- "The World Heritage List: Which Sites Promote the Brand? A Big Data Spatial Econometrics Approach" Journal of Cultural Economics (2016) with Marc Patry
- "What is the Value of World Heritage Status for a German National Park?" Tourism Economics

*Other Publications (Reports, Books Chapters)*

- "Opportunities and constraints to integrate Ghana's smallholder farmers into global value chains" in African Economic Outlook, a joint publication by the African Development Bank, the OECD and the UNDP(2014)

*Selected Conference Presentations and Posters*

- The Role of Self-Efficacy in Behavioral Economics (Annual Conference of the American Agricultural Economics Association 2016) (Pres.)
- Self-Efficacy and Technology Adoption: Understanding Farmer Responses to Rainfall Change (with David Zilberman and Johannes Sauer)(CSAE at Oxford University 2016) (Pres.)
- Experiences in the Gold Coast and Contract Farming in Ghana (with Johannes Sauer)(Annual Conference of the International Association for Applied Econometrics 2015 (Pres.), Annual Conference of the American Agricultural Economics Association 2015 (Poster), and GEWISOLA 2015 (Pres.))
- Sustainable Intensification of Pineapple Farming in Ghana (with Johannes Sauer and Linda Kleeman)( Bioecon Conference at University of Cambridge 2014 (Pres.))
- Estimating a Recreational Demand Model for Jasmund National Park (Conference of the European Association for Environmental and Resource Economists 2013 (Pres.))

*Teaching Experience*

| | |
|---|---|
| S Semester 2016 | Applied Statistics and Econometrics, Undergraduate, TU Munich |
| S Semester 2016 | Climate Change Economics, Graduate, TU Munich |
| W Semester 2016 | Statistical Methods, Graduate, Technical University Munich |
| W Semester 2016 | Applied Econometrics. Graduate, TU Munich |
| S Semester 2015 | Applied Statistics and Econometrics, Undergraduate, TU Munich |
| S Semester 2015 | Climate Change Economics, Graduate, TU Munich |
| W Semester 2014 | Statistical Methods, Graduate, Technical University Munich |
| S Semester 2014 | Climate Change Economics, Graduate, TU Munich |
| S Semester 2012 | Environmental Economics, Undergraduate, University of Kiel |

*Research Visits and Workshops*

| | |
|---|---|
| Summer 2015 and 16 | Research Visit UC Berkeley – Prof. David Zilberman |
| June 2015 | “Mastering `Metrics – From Cause to Effect” (Seminar), Ravello, - Prof. Joshua Angrist |
| June 2013 | “Discrete Choice Analysis: Predicting Demand and Market Shares“ (Seminar) MIT, Cambridge, USA. – Professor Mosche Ben-Akiva |
| Febr. 2013 | “Aspects of Statistical Modelling with Survey Data” (Seminar) Leibnitz-Institute for the Social Sciences, Köln, Germany |
| Nov. 2011-Feb2012 | Research Visit UNESCO World Heritage Center |

*Projects*

- FITHYDRO (H2020)
- Groundwater and Dependent Ecosystems: New Scientific and Technological Basis for Assessing Climate Change and Land-use Impacts on Groundwater (work Package contributor)(EU FP7)
- Culture and Rural Development in Ghana (University of Kiel)
- Evaluating the Bavarian "Ökomodell" Program (Bavarian Ministry of Agriculture)
- The Effects and Value of small and family farms in Europe (ministries of agriculture in Bavaria, Tirol, Austria, and Norway)
- The Economics of UNESCO's World Heritage Sites (UNESCO World Heritage Center and TripAdvisor)