Three Misallocations

by

Ajay Shenoy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2014

Doctoral Committee:

Assistant Professor Raj Arunachalam, Co-Chair
Professor David Lam, Co-Chair
Professor Daniel Ackerberg
Assistant Professor Achyuta Adhvaryu
Associate Professor Christopher House

I dedicate this dissertation to everyone—family, advisors, teachers, and friends—who believed in me. I do not understand how I have earned the faith of such wonderful people, but I am grateful.

ACKNOWLEDGMENTS

Above all I must acknowledge the tireless efforts and inhuman patience of my committee: Raj Arunachalam, David Lam, Dan Ackerberg, Achyuta Adhvaryu, and Chris House.

In addition, I must acknowledge the help I received on each separate chapter:

Acknowledgements for Chapter 1: Special thanks to Raj Arunachalam, Chris House, David Lam, Dan Ackerberg, Ach Adhvaryu, Jagadeesh Sivadasan, Rebecca Thornton, Jeff Smith, Dmitriy Stolyarov, Matthew Shapiro, Miles Kimball, Penelope Goldberg, Isaac Sorkin, Eric Lewis, Ryotaro Tashiro, Jochen Gunter, Lasse Bruno, Adam Larson, four anonymous referees, and many seminar participants for valuable feedback. I also acknowledge generous financial support from the Thai Studies Grant Program at the Center for Southeast Asian Studies at the University of Michgian. I am also grateful to Robert Townsend and the Townsend Thai Project for collecting and distributing the data I use for this project.

Acknowledgements for Chapter 2: Special thanks to Raj Arunachalam, David Lam, Jeff Smith, David Weil, Anja Sautmann, Chris Udry, Silvia Prina, Stefanie Stantcheva, seminar participants at Michigan and Brown, and conference attendees at PacDev and MWIEDC in 2013 for helpful comments and advice. I am also grateful to Robert Townsend and the Townsend Thai Project for collecting and distributing the data I use for this project.

Acknowledgements for Chapter 3: Special thanks to Raj Arunachalam, David Lam, Andrew Foster, Jeff Smith, Jing Cai, Rob Garlick, and the seminar participants at Michigan and Brown for valuable suggestions. I am very grateful to Andrew Foster for the data that made this research possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

Market Failures and Misallocation: Separating the Costs of Factor and Financial Market Failures

1.1 Introduction

Two decades of research in growth have concluded that differences in aggregate resources cannot explain differences in aggregate output, and economists have now asked whether the allocation of aggregate resources can. Hsieh and Klenow (2009), for example, argue India could raise manufacturing output by 40 to 60 percent if it allocated its labor and capital among firms as efficiently as the U.S., and by 100 percent if the allocations were perfect. Yet their study cannot tell where the observed misallocation comes from. Theorists cannot build accurate models and policymakers cannot design useful interventions unless they know what frictions cause misallocation.

I develop a method to measure and separate the misallocation caused by factor and financial market failures, two frictions that set poor countries apart from rich. With perfect markets each firm's optimal allocation depends only on two sets of parameters: its productivity and the common production function. I use a dynamic panel approach to estimate the parameters and calculate the optimal allocation. I define the increase in output from optimal reallocation to be the total misallocation from factor and financial market failures. To separate the cost of each market's failure I exploit how factor markets govern the firm's mix of inputs. When factor markets are perfect, a firm can optimally divide its spending between inputs. By perfecting each firm's mix of inputs while holding its scale constant, I place a lower bound on the aggregate gains from perfecting factor markets, and by then perfecting scale I place an upper bound on the gains from subsequently perfecting financial markets. Finally, I decompose aggregate output into three components: an aggregate production function, average firm-level productivity, and the efficiency of factor allocations. I calculate the counterfactual path of aggregate output if growth in any component were shut down. The counterfactuals show how output would have grown if factor allocations had not improved.

My method only works in a setting where factor and financial market failures cause all misallocation, and I study Thailand's rice sector because it fits the assumption well. Since rice production is relatively uniform I can estimate a common production function and correctly calculate each farmer's marginal product, minimizing the spurious misallocation caused by measurement error. Other sources of misallocation like monopoly, taxes, and adjustment costs are rare in rice farming. It is also plausible that a farmer chooses land, labor, and capital with equal information about her productivity, letting me calculate her ideal mix of inputs, which I use to separate factor and financial market failures.

I find surprisingly little misallocation. The overall cost is 15 percent of output in 1996 and falls to 4 percent by 2008. By then most misallocation comes from factor markets rather than financial markets. Decreases in misallocation contributed little to growth in aggregate rice output relative to growth from factor accumulation and rising average productivity. As Thailand industrializes, labor and capital have left the rice sector. I find that their departure would have caused rice output to fall if farms had not also grown more productive. Finally, I assess the impact on misallocation of the quasi-experimental Million Baht credit program first studied by Kaboski and Townsend (2011). Additional credit has a statistically significant but small effect. A one percent increase in credit reduced misallocation by .1 percentage points, and as expected the program worked almost entirely by reducing misallocation from financial market failures.

Economists started measuring misallocation in response to evidence that financial markets in poor countries do not allocate capital efficiently.[1] Communal divisions inefficiently concentrate capital among incumbent garment manufacturers in India (Banerjee and Munshi, 2004), lending arrangements fail to perfectly insure households in Nigeria (Udry, 1994), and entrepreneurs could reap large returns with small capital investments in Sri Lanka (De Mel et al., 2008). Until recently, however, few studies tried to quantify the aggregate costs of misallocation, prompting Banerjee and Duflo (2005) to emphasize the question and its implications. Jeong and Townsend (2007) explored the idea with a structural model where credit constraints prevent households from buying capital and switching sectors. Their model reproduces much of the change in Thailand's Solow residual even with zero technological progress. With few exceptions (e.g. Benjamin, 1992), the literature

[1] In addition to those I cite in the main text, other recent papers on misallocation include Restuccia and Rogerson (2008); Banerjee and Moll (2010); Peters (2011); Bollard et al. (2012); Alfaro et al. (2008); Bartelsman et al. (2009); Jones (2011); Osotimehin (2011); Alfaro and Chari (2012); Moll (2010).

has focused primarily on financial markets—markets for credit and insurance—and paid little attention to factor markets, where households hire labor and rent land and capital. But if factor markets cause most misallocation, governments ought to worry less about getting credit to the poor and focus on, for example, eliminating regulations that distort the market for land.

Hsieh and Klenow's study (2009), on the other hand, calculates the cost of all misallocation regardless of the source. They build a model of monopolistic competition where firms pay "taxes" to sell output and rent capital. The authors derive the aggregate gains to reallocating factors within manufacturing industries and plug in production functions calibrated with U.S. data. They argue India and China's distorted factor and financial markets are why they have more misallocation than the U.S., but their approach cannot actually link misallocation to its cause. I too compare the actual allocation to more optimal allocations, but my method and context let me attribute misallocation to factor and financial market failures. Midrigan and Xu (2013) approach a question similar to mine with a different method and find a similar answer. Instead of measuring misallocation caused by financial market failures, they calibrate a model of credit-constrained firms and find it cannot predict much of the variation in the marginal product of capital.

To my knowledge this paper is the first to split misallocation into the contributions of factor versus financial markets, and I find earlier work may have been wrong to ignore factor markets. I also measure misallocation under more plausible assumptions than earlier work. For example, since rice production is relatively uniform I can estimate the production function instead of assuming U.S. parameters describe Thailand's production. In addition, I measure misallocation at each point in time using a household survey, making sample attrition less likely to bias the result. Finally, I link misallocation to its sources under weaker assumptions than earlier work, which uses structural models to simulate how much misallocation would occur if each constraint firms face were eased or tightened. Only if the constraints are modeled correctly can they accurately link misallocation to its causes. In contrast, I do not have to assume I perfectly model constraints to measure misallocation from factor versus financial markets.

1.2 Misallocation among Thai Rice Farmers

My crucial assumption is that all misallocation in the Thai rice sector comes from factor or financial market failures. Comparing the original allocation of land, labor, and capital to the allocation that equalizes marginal products does not work if marginal products differ for reasons other than factor and financial market failures. The econometrician might miscalculate marginal products and find misallocation where there is none. He may assume the wrong technology or incorrectly assume firms use the same technology. Unanticipated productivity shocks might change firms' marginal products after they choose their factors, making the allocation look inefficient even when markets are perfect. Real forces other than weak factor and financial markets might also drive marginal products apart. Firms may pay different taxes, have adjustment costs, or be monopolists.

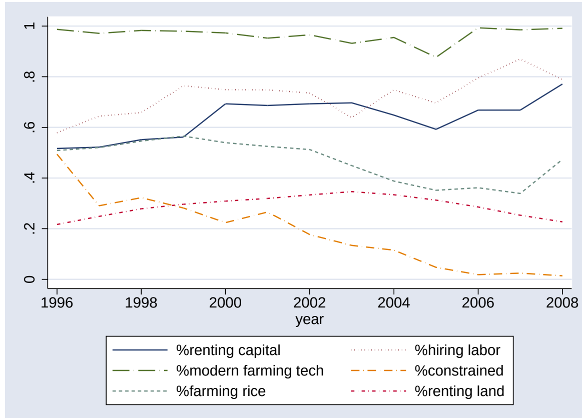
Such issues cause fewer problems in rice farming than in manufacturing. Rice production is relatively uniform and transparent. Though not all farmers in Thailand grow the same type of rice, they grow each variety following a similar technique.[2] And unlike in many developing countries nearly all farmers use modern pesticides and fertilizers. Figure 1.1 shows that nearly 100 percent of my sample uses modern farming technology (fertilizers or pesticides) throughout the entire sample period. [3] I can estimate a common production function for all farmers without departing too far from the truth. Since identifying the sources of anticipated versus unanticipated productivity is easier in rice production—a farmer knows his own talent but does not know whether rats will eat his crop after the harvest—I can model productivity as described in Section 1.5.2.

Monopoly, taxation, and adjustment costs are not big problems in the rice sector, either. Rice is a commodity and Thai farmers are all price-takers who sell their output to mills and merchants at market prices. In Shenoy (2014b) I show that farmers' selling prices move with the international rice price. Though the government often supports prices, price subsidies will affect all farmers equally and leave allocations unchanged. Thailand's farmers grow rice more commercially than their Indian or Chinese counterparts, but they enjoy a similar lack of taxes. Of the roughly 1500 survey households who reported any agricultural activity in 1996, only eight reported paying land taxes. Less than two percent of rice farmers in the monthly survey report paying any income tax. My assumption that farmers in a village can exchange factors without adjustment costs is also plausible. Tractors have wheels and bullocks have legs, most on-farm machinery is not too large to move, and one farmer can store his crops in another's granary. There is no cost to hiring or firing a casual farm worker. Exchanging land is not difficult, as over three quarters of the rice paddies

[2]The farmer seeds a nursery plot and transplants the seedlings to a flooded paddy where they grow to adulthood. Farmers fertilize and apply pesticides until the rice matures and they harvest, thresh, dry, and sell the grains. According to the International Rice Research Institute, most of Thailand's farmers use this lowland rain-fed method to grow their rice.

[3] Thailand was an early ally of the U.S. during the Cold War and received American aid to modernize its rice sector in the 1960s. Despite fears to the contrary, both small and large farmers adopted the new seeds and fertilizers. The adoption of fertilizer was so rapid that, according to Baker and Phongpaichit (2009), a Japanese anthropologist visiting in 1970 found "Villagers who had described the local rituals to him only a decade ago now exclaimed 'the rice spirit is no match for chemical fertilizer'."

Figure 1.1
Characteristics of Thai Rice Farmers



Note: Descriptive statistics of the sample. I describe the sample in more depth in Section 1.4.

cultivated in 1996-1997 are no more than two kilometers from the village.

Not only are other sources of misallocation absent, but also the assumptions I make to calculate the optimal allocation are more plausible among Thai rice farmers. Two of those assumptions are that productivity is Hicks-Neutral and technology is Cobb-Douglas. Productivity need not be Hicks-Neutral—that is, affect all factors of production equally—in manufacturing or services. If demand for sneakers rises, Nike's marketing team does more to increase its sales than does the quality of its machinery. Different factors at Nike essentially make different products. Machines and unskilled labor make shoes while skilled workers make a marketing campaign. In rice farming, however, all factors work to make a single product. The farmer's year-to-year productivity shocks—rainfall and pests—will damage the end crop, not the individual contribution of the workers or the tractors. If the farmer falls ill the lost productivity comes from a reduction in her managerial ability, as she can hire others to supply labor. Major advances like the invention of the tractor are unlikely in my thirteen-year sample. The Cobb-Douglas assumption—that the optimal share of expenditure on labor versus capital does not change with scale—is also more plausible because the farmer will farm each additional plot of land much like the previous. I also verify the assumption approximates reality in Appendix A.5.[4]

To split factor market and financial market misallocation I approximate what would happen if each village perfected its factor markets but left its financial markets untouched.

[4]The general model and procedure I present in Appendix A.3 does not require any assumptions beyond concavity, decreasing returns, and twice-differentiability. But estimating a more complicated production function requires stronger assumptions about productivity and factor input choices.

Assuming (as I have argued) adjustment costs are small, with perfect factor markets a farmer will always choose the optimal mix of factors regardless of overall expenditure. To calculate the optimal mix I assume the farmer chooses all factors with the same information about within-village productivity shocks. The assumption, which implies all farmers choose the same mix, would be too strong if it applied to all shocks. A farmer does not know how much rain will fall when he chooses how much land to sow, but he does when he hires workers at harvest. Rainfall and other village-level shocks, however, affect all farmers in the village and have no allocative effects. Most within-village shocks, such as when pests eat freshly harvested rice, happen after all factors are chosen or show no effects until harvest. Shocks like illness, however, can strike after renting land but before hiring labor or renting tractors. If such shocks are common I may incorrectly measure the misallocation from factor market imperfections. But illness is not common (see Section 1.4), and I show in Section 1.7 that my measures of factor and financial market misallocation respond as expected to a credit intervention. [5]

A common problem in panels of manufacturing firms is that unproductive firms leave the industry and equally unproductive ones enter, but the entrants do not appear in the panel. Over time the sample becomes more productive than the population. Since households still farm most of Thailand's rice, however, I can measure production using a household survey, which follows both farmers and non-farmers. New entrants to farming—households who did not farm at baseline but start farming later on—will enter my sample of farmers at the same rate as the population because the survey was representative at baseline. Migration, which entirely removes households from the survey, is relatively rare. Since the Project sampled new households to replace the ones who left, the sample does not become too unrepresentative. I do not need to assume a stable sample of farmers because exit does not bias my estimate of misallocation. When a farmer exits, he takes his factors out of rice production. The farmers and factors that remain represent the sector's current state, and misallocation is still the output gained from redistributing the factors more efficiently.[6]

[5] What if farmers can get credit to purchase some factors but not others? The most obvious example is store credit for seeds and fertilizer. With perfect factor markets a farmer can arbitrage borrowed seeds and fertilizer to acquire land or capital, so my argument about perfect mix still holds. Regardless, few households report borrowing anything on store credit at baseline.

[6] I must exclude from the sample farmers who farm only once because I cannot calculate a fixed effect for them. If they are severely over- or under-allocated, villages will falsely appear efficient. A one-time farmer, however, can only enter and exit if the rental or purchase markets work well. As villages with efficient markets have little misallocation, if anything the bias works towards finding too much misallocation because the efficient villages receive less weight in the sector-level calculation of efficiency.

1.3 Measuring Misallocation

I now present a model of household rice production and derive from it my measures of misallocation. In building the model I sacrifice generality in favor of clarity and consistency. The model reflects rice production in Thailand as described in Section 1.2 and imposes the functional forms I use to compute my measures in Section 1.6. I describe a more general approach in Appendix A.3.

1.3.1 Environment

The farmer maximizes her discounted lifetime utility from consumption. She uses capital K, land T, and labor L to produce farm revenue y. Let $\boldsymbol{X}$ denote the vector of productive factors chosen. The factors may come from her stock of owned factors $\boldsymbol{X}^o$ or the factors she rents from factor markets $\boldsymbol{X} - \boldsymbol{X}^o$. She buys $\boldsymbol{I}_{it}$ factors, and owned factors depreciate at rates recorded in the matrix $\boldsymbol{\delta}$. The vector of "owned" factors $\boldsymbol{X}^o$ includes family labor, which I assume exogenous for notational simplicity (making labor decisions endogenous changes nothing because I only care about labor allocations between farms). I implicitly assume family labor and hired labor are perfect substitutes, though relaxing the assumption does not change the result much.[7] The farmer buys factors (except labor) at a vector of prices $\boldsymbol{p}$ and rents them at prices $\boldsymbol{w}$. Her output also depends on Hicks-Neutral productivity, part of which she anticipates when choosing factors (A) while the rest (ϕ) is random and unanticipated.

Farmer i faces her own set of prices, a budget constraint, a liquidity constraint, and a factor market constraint. At the beginning of each year she spends z on renting and buying factors. Any spending beyond gross interest on her savings from last year R^b is borrowed at gross rate R^z and repaid after the harvest. She faces a liquidity constraint $\bar{z}$, which is a function of the owned land she offers as collateral and differs across farmers (perhaps because they differ in how many cosigners or rich relatives they have). She cannot rent in more factors than $\overline{\boldsymbol{X}}$ or rent out more factors than $\underline{\boldsymbol{X}}$. Since she lacks perfect insurance her consumption may be correlated with the unanticipated shock ϕ, and the effect on her utility depends on the parameters $\boldsymbol{\gamma_i}$. The Lagrange multipliers λ, ω, and $\underline{\kappa}, \overline{\kappa}$ capture the shadow costs of the budget, liquidity, and factor market constraints.

Formally, the farmer's problem is

$$\text{Maximize} \qquad \mathbb{E}[\sum_{j=0}^{\infty} \rho^j u(c_{i,t+j}; \boldsymbol{\gamma_i})]$$

[7] The sensitivity analysis is available upon request.

Subject to:

$$
\begin{aligned}
& y_{it} = A_{it}\phi_{it}K_{it}^{\theta_K}T_{it}^{\theta_T}L_{it}^{\theta_L} && (Production\ Function) \\
& z_{it} = \boldsymbol{w}_{it}(\boldsymbol{X}_{it} - \boldsymbol{X}_{it}^o) + \boldsymbol{p}_{it}\boldsymbol{I}_{it} && (Expenditure) \\
& \boldsymbol{X}_{it}^o = \boldsymbol{I}_{it} + (1-\boldsymbol{\delta})\boldsymbol{X}_{i,t-1}^o && (Owned\ Assets) \\
\lambda: \quad & c_{it} + b_{i,t+1} = y_{it} - R_{it}^z(T_{it}^o)(z_{it} - R_{it}^b b_{it}) && (Budget\ Const.) \\
\omega: \quad & z_{it} - R_{it}^b b_{it} \le \bar{z}_{it} && (Liquidity\ Const.) \\
\underline{\kappa}, \overline{\kappa}: \quad & \underline{\boldsymbol{X}}_{it} \le \boldsymbol{X}_{it} - \boldsymbol{X}_{it}^o \le \overline{\boldsymbol{X}}_{it} && (Factor\ Market\ Const.)
\end{aligned}
$$

1.3.2 Perfect Choices and Distortions

The farmer's optimal choice of capital satisfies

$$\mathbb{E}[\lambda_{it}\phi_{it}]\theta_K A_{it}K_{it}^{-(1-\theta_K)}T_{it}^{\theta_T}L_{it}^{\theta_L} - (R_{it}^z(T_{it}^o)\mathbb{E}[\lambda_{it}] - \omega_{it})w_{it}^K + \underline{\kappa} - \overline{\kappa} = 0$$

Optimal land and labor choices satisfy similar conditions. When markets are fully perfect the unanticipated shock does not affect consumption ($\mathbb{E}[\lambda_{it}\phi_{it}] = \mathbb{E}[\lambda_{it}]\mathbb{E}[\phi_{it}]$), farmers pay the same borrowing rate ($R_{it}^z[T_{it}^o] = R_t^z$), the liquidity constraint does not bind ($\omega_{it} = 0$), farmers pay the same rental prices ($w_{it}^K = w_t^K$), and factor market constraints do not bind ($\underline{\kappa}_{it} = \overline{\kappa}_{it} = 0$). Then

$$
\begin{aligned}
\theta_K A_{it}K_{it}^{-(1-\theta_K)}T_{it}^{\theta_T}L_{it}^{\theta_L} &= R_t^z w_t^K \\
&= \theta_K A_{jt}K_{jt}^{-(1-\theta_K)}T_{jt}^{\theta_T}L_{jt}^{\theta_L} \quad \forall j
\end{aligned} \tag{1.1}
$$

The expression implies marginal products are equalized across all farmers in the village.

Now suppose factor markets are perfect, so $w_{it}^K = w_t^K, \underline{\kappa}_{it} = \overline{\kappa}_{it} = 0$, but financial markets are not. Then the optimal choices of capital and land satisfy

$$
\begin{aligned}
\mathbb{E}[\lambda_{it}\phi_{it}]\frac{\mathbb{E}[y_{it}]}{K_{it}} &= \frac{1}{\theta_K}(R_{it}^z(T_{it}^o)\mathbb{E}[\lambda_{it}] - \omega_{it})w_t^K \\
\mathbb{E}[\lambda_{it}\phi_{it}]\frac{\mathbb{E}[y_{it}]}{T_{it}} &= \frac{1}{\theta_T}(R_{it}^z(T_{it}^o)\mathbb{E}[\lambda_{it}] - \omega_{it})w_t^T.
\end{aligned}
$$

Divide the capital condition by the land condition:

$$\begin{aligned}\frac{T_{it}}{K_{it}} &= \frac{\theta_T}{\theta_K}\frac{w_t^K}{w_t^T} \\ &= \frac{T_{jt}}{K_{jt}} \quad \forall j\end{aligned} \tag{1.2}$$

The condition implies capital-land ratios are equalized across farmers throughout the village, and by similar logic the land-labor ratios are as well.

1.3.3 Optimal Allocations

How would perfect markets allocate the aggregate factor stock? Suppress time subscripts for notational simplicity and suppose i is a farmer in village I observed to use $\bar{K}_i, \bar{T}_i, \bar{L}_i$. [8] The observed factor stocks are $K_I = \sum_{i \in I} \bar{K}_i$ and so on, and they do not change because I only reallocate the village's existing resources. With aggregate stocks pinned down, I can ignore the supply side of the market and normalize $R^z = 1$. Use (1.2) to eliminate T_i and L_i from (1.1):

$$K_i^* = \left[A_i \left(\frac{\theta_K}{w^K}\right)^{1-\theta_T-\theta_L} \left(\frac{\theta_T}{w^T}\right)^{\theta_T} \left(\frac{\theta_L}{w^L}\right)^{\theta_L}\right]^{\frac{1}{1-\sigma}}$$

Since the aggregate stock does not change, the market-clearing condition is

$$\begin{aligned}K_I &= \sum_{j \in I} K_j^* \\ &= \left[\left(\frac{\theta_K}{w^K}\right)^{1-\theta_T-\theta_L} \left(\frac{\theta_T}{w^T}\right)^{\theta_T} \left(\frac{\theta_L}{w^L}\right)^{\theta_L}\right]^{\frac{1}{1-\sigma}} \sum_{j \in I} A_j^{\frac{1}{1-\sigma}} \\ \Rightarrow \frac{K_I}{\sum_{j \in I} A_j^{\frac{1}{1-\sigma}}} &= \left[\left(\frac{\theta_K}{w^K}\right)^{1-\theta_T-\theta_L} \left(\frac{\theta_T}{w^T}\right)^{\theta_T} \left(\frac{\theta_L}{w^L}\right)^{\theta_L}\right]^{\frac{1}{1-\sigma}}\end{aligned}$$

Plug this back into the individual factor choice to derive the optimal allocations with fully perfect markets:

[8] Since factor markets must clear at any point in time and the conditions for perfect choices (1.1 and 1.2) do not involve the past or future, dynamics are irrelevant from here onwards.

$$K_i^* = \frac{A_i^{\frac{1}{1-\sigma}}}{\sum_{j\in I} A_j^{\frac{1}{1-\sigma}}} K_I \tag{1.3}$$

Optimal land and labor are similar. Call farmer i's output with perfect allocations $y_i^* = A_i\phi_i(K_i^*)^{\theta_K}(T_i^*)^{\theta_T}(L_i^*)^{\theta_L}$.

Now suppose factor markets are perfected but financial markets left untouched. Equation 1.2 gives the optimal mix but not the overall scale of factors for each farmer, so I impose that farmers' purchases in the perfected factor markets add up to the value of their endowments:

$$w^K K_i^+ + w^T T_i^+ + w^L L_i^+ = w^K \bar{K}_i + w^T \bar{T}_i + w^L \bar{L}_i$$

where K_i^+ is the farmer's new choice of capital while $\bar{K}$ is still her original choice, now treated like an endowment. Any assumption about scale would define the allocation, but this one is appealing because it casts the farmers into an Edgeworth economy where their original factor choices are like endowments. The farmer choosing a profit-maximizing mix of factors behaves like a consumer choosing a utility-maximizing bundle of goods. The resulting allocation is easy to compute and perfects each farmer's mix of factors while leaving her scale fixed. Each farmer's allocation may differ from what she would choose given perfect factor markets and no extra constraints on scale. But under an assumption I explain in Section 1.3.4, gains from moving to the computed allocation are a lower bound on true misallocation from imperfect factor markets.

The farmer spends a constant fraction of her wealth on each factor, so the condition for capital is

$$K_i^+ = \frac{\theta_K}{\theta_K+\theta_T+\theta_L}\cdot\frac{w^K \bar{K}_i + w^T \bar{T}_i + w^L \bar{L}_i}{w^K}.$$

The aggregate factor stock does not change, so markets clear as follows:

$$\begin{aligned}
K_I &= \sum_{j\in I} K_j^+ \\
&= \frac{\theta_K}{\theta_K+\theta_T+\theta_L}\frac{1}{w^K}\sum_{j\in I}(w^K \bar{K}_j + w^T \bar{T}_j + w^L \bar{L}_j) \\
&= \frac{\theta_K}{\theta_K+\theta_T+\theta_L}\frac{1}{w^K}\left(w^K K_I + w^T T_I + w^L L_I\right).
\end{aligned}$$

Divide by the analogous condition for land to derive an aggregate version of (1.2):

$$\frac{K_I}{T_I} = \frac{\theta_K}{\theta_T}\frac{w^T}{w^K}.$$

Substitute this back into the farmer's optimal choice to derive the allocations under perfect factor markets:

$$\begin{aligned} K_i^+ &= \frac{\theta_K}{\theta_K+\theta_T+\theta_L} \cdot \left[\frac{w^K}{w^K}\bar{K}_i + \frac{w^T}{w^K}\bar{T}_i + \frac{w^L}{w^K}\bar{L}_i\right] \\ &= \frac{1}{\theta_K+\theta_T+\theta_L}\left[\theta_K\frac{\bar{K}_i}{K_I} + \theta_T\frac{\bar{T}_i}{T_I} + \theta_L\frac{\bar{L}_i}{L_I}\right]K_I \end{aligned} \tag{1.4}$$

Optimal land and labor are similar. Call farmer i's output with perfect factor markets $y_i^+ = A_i\phi_i(K_i^+)^{\theta_K}(T_i^+)^{\theta_T}(L_i^+)^{\theta_L}$.

1.3.4 Costs of Misallocation

For all three scenarios, aggregate output in village I is the sum of each farmer's output: actual output Y_I, output with fully perfect markets Y_I^*, and output with only perfect factor markets Y_I^+. I use two measures of misallocation: the gains from perfecting one or more markets, and the fraction of efficient output achieved. The gains from reallocation (or simply "misallocation") measure how much output a village loses from misallocations. The fraction of efficient output achieved (or "efficiency") compares the real world to the world with perfect markets and appears naturally in the aggregate production function I derive in Section 1.3.5. Define

$$\begin{aligned} G_I &= \frac{Y_I^* - Y_I}{Y_I} \qquad & G_I^{FACT} &= \frac{Y_I^+ - Y_I}{Y_I} \qquad & G_I^{FIN} &= \frac{Y_I^* - Y_I^+}{Y_I} \\ E_I &= \frac{Y_I}{Y_I^*} & E_I^{FACT} &= \frac{Y_I}{Y_I^+} & E_I^{FIN} &= \frac{Y_I^+}{Y_I^*}. \end{aligned}$$

The gains from perfecting each market add up to the overall gains ($G_I = G_I^{FACT} + G_I^{FIN}$), and overall efficiency is the product of factor and financial market efficiency ($E_I = E_I^{FACT} \cdot E_I^{FIN}$). The overall gains are a decreasing function of efficiency ($G_I = \frac{1}{E_I} - 1$).

The factor market misallocation I compute (G_I^{FACT}) need not equal the true gains from perfecting factor markets. I compute factor market misallocation by holding each farmer's scale of production fixed, but if factor markets actually became perfect a productive farmer would probably increase her scale. Since she can allocate each dollar to the factor she needs most, every dollar spent on rice farming yields more revenue than it would with imperfect factor markets. Proposition 1, which I prove in Appendix A.1, formalizes this argument:

Proposition 1 *Let $\tilde{K}_i^+$ be the level of capital farmer i would choose if factor markets were perfected, financial markets left untouched, and the endowment constraint were not imposed. Assume $\mathbb{E}[A_i((\tilde{K}_i^+)^\sigma - (K_i^+)^\sigma)] > 0$. Then in expectation G_I^{FACT} is a lower bound on the true gains from perfecting factor markets and G_I^{FIN} an upper bound on the true gains from subsequently perfecting financial markets.*

The assumption states that with perfect factor markets the most productive farmers will increase their scale relative to the actual outcome. [9] The only reason the assumption might fail is if factor market failures somehow compensate for financial market failures, such as if the farmers who cannot get bank loans can rent land more cheaply than everyone else, and they are also the most productive farmers. The scenario is implausible in a poor rural village, where those shut out of financial markets are usually shut out of factor markets as well.[10]

1.3.5 Decomposing Aggregate Output and Growth

Growth accounting traditionally measures changes in per capita output and not per firm output, so I decompose the growth in the village's rice output per household instead of per farmer. Suppose $\mathcal{I}$ is the set of all households (rice-farming or otherwise) in village I, and let $\mathcal{Y} = \frac{Y}{|\mathcal{I}|}$ be per household rice output. Let $Z_{it} = A_{it}\phi_{it}$ denote overall productivity, Z_{It} its mean and $\tilde{Z}_{it}$ deviations from the mean. I use overall productivity to be consistent with traditional growth accounting (which computes an overall Solow residual) and because it is difficult to split aggregate shocks into anticipated and unanticipated parts. [11]

[9]Capital simply stands in for scale of production. I could have phrased the proposition in terms of land or labor just as easily because the ratios of all factors are fixed by (1.2).

[10]The proof is an equivalency result, so if the assumption failed and $\mathbb{E}[A_i((\tilde{K}_i^+)^\sigma - (K_i^+)^\sigma)] < 0$, the computed gains would be an upper bound. In the knife-edge case where $\mathbb{E}[A_i((\tilde{K}_i^+)^\sigma - (K_i^+)^\sigma)] = 0$, computed gains equal actual gains.

[11]For example, how much of a district-year dummy is anticipated? The distinction does not matter for within-village reallocation but will affect how much growth is assigned to anticipated versus unanticipated aggregate productivity. Rather than make an arbitrary and misleading distinction I combine the two and call the result overall productivity.

Then

$$
\begin{aligned}
\mathcal{Y}_{It} &= Z_{It}E_{It} \cdot \frac{1}{|\mathcal{I}|}\sum_{i\in I}\tilde{Z}_{it}(K^*_{it})^{\theta_K}(T^*_{it})^{\theta_T}(L^*_{it})^{\theta_L} \\
&= Z_{It}E_{It}F(K_{It}, T_{It}, L_{It}; \{\tilde{A}_{it}\}\{\tilde{\phi}_{it}\})
\end{aligned}
$$

Recall from (1.3) the optimal factor allocations $K^*_{it}, T^*_{it}, L^*_{it}$ are only functions of the aggregate stocks and relative productivity. So taking the relative productivity distribution $\{\tilde{A}_{it}\}, \{\tilde{\phi}_{it}\}$ as a parameter, F is a function of only aggregate capital, land, and labor—the aggregate production function.

I can derive a similar decomposition for sample-wide output $\mathcal{Y}_t$, but since households were sampled into the survey in multiple stages the decomposition must weight villages by their size. Let κ_{It} be the population of village I as a fraction of the total population of all villages surveyed, and let $\chi_{It} = \frac{\kappa_{It}\mathcal{Y}^*_{It}}{\sum_I \kappa_{It}\mathcal{Y}^*_{It}}$ be the share of sample-wide output it produces under optimal within-village allocations. Let $Z_t = \sum_I \chi_{It}Z_{It}$ denote the output-weighted mean and $\tilde{Z}_{It}$ deviations from the mean of village productivity. Let E_t be sample-wide allocative efficiency with reallocation still within villages. Trivial algebra shows $E_t = \sum_I \chi_{It}E_{It}$, which means

$$
\begin{aligned}
\mathcal{Y}_t &= Z_t \cdot E_t \cdot \sum_I \tilde{Z}_{It}\kappa_{It}F(K_{It}, T_{It}, L_{It}) \\
&= Z_tE_tF(\{K_{It}, T_{It}, L_{It}\}) \qquad (1.5)
\end{aligned}
$$

Since reallocation is within-village, the sample-wide aggregate production function depends on the aggregate factor stocks of each village.

Define g^V_t as the log change of any variable V over baseline. Since $g^{\mathcal{Y}}_t = g^Z_t + g^E_t + g^F_t$, I can decompose growth in per household rice output into the contributions of improvements in productivity, improvements in factor allocations, and aggregate factor accumulation. By setting g^Z_t and g^E_t to zero, for example, I can examine how output would have grown if the rice sector had made no improvements to productivity or efficiency.

1.4 Data

I construct my sample from the Townsend Thai Annual Household Survey (1997). The Townsend Thai Project collected a baseline of households from four rural provinces using a stratified design. The Project subsampled one third of the survey villages and resurveyed the sampled households every following year to construct a panel. It later added two more provinces and sampled new households to counter attrition. I use the 1997 through 2009 rounds. The survey response period is June of the previous year through May, so I label the period covered by the 1997 survey as 1996. The Project followed sixteen of the villages excluded from the annual survey to collect the Townsend Thai Monthly Household Survey (2012). I use the first two years of the monthly survey throughout the paper to confirm facts not found in the annual survey. I use district-level precipitation data computed from the University of Delaware Climactic Project and NASA's Tropical Rainfall Measuring Mission. I also test my micro measures of productivity in the Online Appendix using agro-climactic suitability data from the FAO.

Land is the number of rai (6.25 rai = 1 acre) of paddy the household cultivated (whether owned or otherwise). Labor is the sum of hired and family labor in days worked. Hired labor is the household's expenditures on farm workers divided by the median daily wage in the village. Using the median wage is not ideal, but the survey does not ask directly about the amount of labor hired and the within-village variation in unskilled wages is relatively low (the coefficient of variation is less than 0.19 for most village-years). I count the number of household members who report being unpaid family laborers with primary occupations in farming of any sort (or who mention "FIELDS" in a freeform response). The annual survey gives no information on the days each member worked. Instead I use the more detailed labor data in the monthly survey to calculate the median days any individual works on his family's fields, and multiply the overall median—60 days—by the number of family laborers counted in the annual data.

Capital is the sum of the value of owned mechanical capital, the value of owned buffalo, and the imputed value of rented capital and expenses (including intermediate inputs). I do not compute the value of owned capital using perpetual inventory because households frequently report changes in capital ownership without reporting investment, either because they forget to report the investment or because they receive the new machines as gifts or inheritance. Instead I assign a purchase value to each asset the household owns. I deflate and depreciate the purchase value of assets owned at baseline. For assets acquired afterwards I use the purchase price. The survey only reports assets in classes, so if the household has multiple assets of the same type I must treat them as if they have identical value and use the most recent purchase price (most households own one or fewer assets of any type). If I cannot identify a price I drop the asset from the calculation (I can identify a price for the large majority). I then depreciate the purchase price to get the value in a given year assuming 2 percent depreciation for structures (House and Shapiro, 2008), 10 percent depreciation for machines, and (I treat them as vehicles) 20 percent depreciation for tractors

(Levinsohn and Petrin, 2003). Owned mechanical capital in a year is the total value of the assets. I treat intermediate inputs—seeds, fertilizers, pesticides, and fuel—as capital with a 100 percent deprecation rate. I add maintenance, which I treat as investment that takes immediate effect. I then add the purchase price of rented capital, which I approximate with total rental expenses divided by an interest rate of .04 plus the average rate of deprecation for all types of capital (a user cost). [12] Finally, I add the value the household reports for its buffalo. Since households do not report whether they rented out their capital I cannot lower it to reflect how much they actually use. The error might inflate estimated misallocation because unproductive farmers who rent out their machinery will appear to have too much capital.[13]

In Section 1.5.2 I model productivity using several catastrophes the household reports about its income. I use indicators for illness, death in the family, flooding, problems with crop-eating pests, poor rainfall, low yield for other reasons, and a low price for output. To proxy for malnourishment I use the share of the household's consumption budget devoted to rice, the staple food, including the value of home-produced rice. Jensen and Miller (2010) argue as households become less hungry they substitute away from the staple, so a larger share implies more hunger. All monetary variables are deflated to 2005 Thai baht. I describe all the variables in more detail in Appendix A.6.

Table 1.1 reports household-year averages of each variable for the sample. I restrict my analysis to the households I observe with positive rice revenue and levels of all factors for at least two years, as I cannot calculate a household fixed-effect for anyone else. At 2005 exchange rates the average annual revenue from rice was roughly 1200 dollars. Farms are small and most farmers plant only 19.4 rai (3.1 acres) of paddy. The average household spent only 40 percent of its annual budget on rice, a sign that hunger is not widespread in my sample. Very few households report illness or deaths in the family relative to the number who report aggregate (rainfall, floods, prices) or completely unanticipated (pests) production shocks. It is reassuring that illness, the type of shock that might undermine my decomposition (see Section 1.2), is rare.

[12]Given the presence of misallocation, how can I assume a common interest rate and a user cost? Recall my objective is to construct a consistent measure of capital in terms of value. Reweighting using household-specific interest rates would be equivalent to calling a tractor more valuable because the household renting it pays a higher mortgage. Variation in the actual rental price is more problematic, but a relatively small fraction of total capital is rented, making this less of a problem.

[13]The survey does ask households about "Payments for other rentals," but the rented goods might not be capital and almost no one reports receiving any (only three do in the latest year when we would expect the best rental markets).

Table 1.1
Sample Descriptives

Revenue from rice	47918.6	Fraction who report. ..	
	(77327.9)		
Capital	74560.7	Illness	0.01
	(93826.4)	Death in Family	0.01
Land	19.4	Flood	0.06
	(15.9)	Crop-Eating Pests	0.04
Labor	181.6	Bad Rainfall	0.22
	(191.9)	Low Yield for Other Reason	0.21
Rice Budget Share	0.4	Low Price	0.15
	(0.2)		
Households	775	Observations	6230
Villages	69		

Note: I construct the sample in two steps: I first restrict it to household-years with positive revenue and positive capital, land, and labor. I further restrict the sample to household with at least two years of positive revenue and factors. All variables are annual. Revenue and capital are in 2005 baht, land in rai, and labor in human-days. The share of the household's consumption budget spent on rice is my measure of hunger.

1.5 Estimating the Production Function

I cannot use the expressions for optimal land, labor, and capital derived in Section 1.3 without an estimate of the production function. Estimating a production function is never easy, and misallocations complicate the task because I cannot assume firms choose their inputs optimally. As I explain, however, misallocation enables me to separately identify each factor's production share using a dynamic panel estimator.

1.5.1 Estimation with Misallocation

I assume in Section 1.2 that farmers make all their input choices with the same information about within-village productivity. Ackerberg et al. (2006) explain why the assumption might make identifying each factor's production elasticity difficult. If a firm using Cobb-Douglas technology selects all its inputs freely and simultaneously they will all be functions of a single variable: anticipated productivity. Equation 1.3 shows that the farmer's optimal choices of (log) capital, land, and labor are all perfectly correlated. Ackerberg et al. solve the problem by assuming the firm chooses factors optimally but at different times.

With imperfect markets the solution is simpler: the firm does not choose factors optimally. If one farmer can rent tractors more cheaply than his neighbor they will use land and capital in different proportions. Even relatively little variation (or misallocation) breaks the perfect correlation between factors. To test whether markets are perfect, take Expression

Table 1.2
Test of Imperfect Markets

	(1) $\frac{Capital}{Labor}$	(2) $\frac{Capital}{Land}$	(3) $\frac{Land}{Labor}$
Capital	0.006***	0.089***	-0.000
	(0.00)	(0.02)	(0.00)
Land	-5.504***	-299.477***	0.004***
	(1.60)	(45.94)	(0.00)
Labor	-1.411**	-1.080	-0.000
	(0.44)	(2.21)	(0.00)
Observations	6230	6230	6230
Households	775	775	775

Note: Under perfect markets the ratios of factor choices should not be correlated with anything after controlling for village-year fixed-effects. I report the results of such regressions. Standard errors are clustered at the household level.

1.2 and add a term for measurement error:

$$\frac{T_{it}}{K_{it}} = \frac{\theta_T}{\theta_K}\frac{w_t^K}{w_t^T} + \varepsilon_{it}$$
$$= v_{It} + \varepsilon_{it}$$

The land-capital ratio is a function of invariant parameters, noise, and prices that do not vary within a village with perfect factor markets. If markets are perfect it should not be correlated with anything—in particular, the levels of any factor—after controlling for village-year fixed-effects. Table 1.2 shows that the capital-labor, capital-land, and land-labor ratios are each strongly correlated with the level of at least one factor. I reject the null in five of nine coefficients, which means I can reject perfect markets. Although the markets cause little misallocation (see Section 1.6), they are imperfect enough to let me identify the production function.

Imperfect markets do, however, rule out a class of methods Ackerberg et al. (2006) call "structural techniques." The econometrician assumes the firm chooses intermediate inputs to match its anticipated productivity. All else equal, more intermediate inputs imply higher anticipated productivity. The assumption yields powerful results when satisfied: the econometrician can non-parametrically estimate productivity. It is also why structural techniques fail when factor markets fail. Take a concrete example and suppose two farmers have the same amount of land, workers, machinery, and productivity. One is married to the son of the local fertilizer merchant. She gets a discount on fertilizer and naturally buys more. The econometrician wrongly concludes she is more productive than her clone. The structural techniques fail because they assume the farmer's choices are only a function of her productivity and not her constraints or privileges.[14]

1.5.2 Modeling Productivity

What makes one farmer more productive than another? Much of what determines a firm's revenue productivity in manufacturing or services—a successful marketing campaign, a new product line, the monopoly power born of a competitor's demise—are absent in agriculture. Many of the most obvious determinants of a farmer's productivity—rainfall, crop-eating pests, illness, accidental misapplication of fertilizer—either affect everyone in the village or are unanticipated. As I argued in Section 1.2, Thai farmers have used modern seeds, pesticides, and fertilizers for decades, so nobody has a technological edge. Land quality does not appear to vary much within a village.[15] Malnourishment might lower the farmer's productivity, and he certainly knows when he is hungry, so I construct a measure of hunger. What remains is the farmer's own managerial talent: his knowledge of how to eke the most output from his inputs of land, labor, and capital. Managerial talent is a fixed characteristic of the farmer I can capture in a household fixed-effect. I model anticipated and unanticipated productivity as follow:

$$\log A_{it} = [Household\ Fixed\ Effect]_i + a^H [Hunger]_{it} \tag{1.6}$$

$$\begin{aligned} \log \phi_{it} = & \sum_j a_j^S [Dummy\ Shocks]_{j,it} + \sum_k a_k^D [\textit{District-Year}\ Dummies]_{k,it} \\ & + \sum_m a_m^R [Monthly\ Precipitation]_{m,it} + [Overall\ Error]_{it} \end{aligned} \tag{1.7}$$

In the dummy shocks I include indicators for illness, death in the family, retirement, flooding, problems with crop-eating pests, poor rainfall, low yield for other reasons, and a low price for output. [16] Why do I call all aggregate shocks and (district-level) precipitation unanticipated? For village-level reallocation the distinction does not matter: a village-level shock will not affect the optimal allocation. But the distinction would change how much growth I credit to anticipated versus unanticipated productivity in the aggregate output

[14] Ackerberg et al. would say the "scalar unobservable" assumption fails.

[15] I took the average price per rai as proxy for land quality and included it in unreported production function estimates. The coefficient was small and insignificant, which I take to mean quality does not vary much and price reflects location rather than fertility.

[16] Self-reports of bad rainfall and prices probably do not add much information to the district-level variables for precipitation and shocks, but I include them anyways to ensure I correctly estimate the production function.

decomposition of Section 1.3.5. I therefore combine both types of productivity into an overall term in the decomposition.

The reader may doubt that I can ever know as much about the farmer's productivity as the farmer himself. In Section 1.8 I assess how much the main results change when farmers anticipate all productivity.

1.5.3 Dynamic Panel Estimation

If the bulk of a farmer's anticipated productivity is fixed, it seems natural to estimate

$$\log y_{it} = \log A_{it} + \tilde{\log}\,\phi_{it} + \theta_K \log K_{it} + \theta_T \log T_{it} + \theta_L \log L_{it} + [Overall\ Error]_{it}$$

with the within-household estimator, where $\tilde{\log}\,\phi_{it} = \log \phi_{it} - [Overall\ Error]_{it}$. But the key assumption for its consistency—what Wooldridge (2002) calls strict exogeneity—fails. Strict exogeneity requires that unexpectedly high or low output in either the past or future will not affect a farmer's input decisions today. But suppose a credit-constrained farmer suffered a bad harvest last year and spent her savings on food, leaving less money to rent land this year. Aside from potentially causing misallocations the situation also violates strict exogeneity.

The Anderson-Hsiao estimator (Anderson and Hsiao, 1981, 1982) can estimate the production function under a weaker assumption called sequential exogeneity. Sequential exogeneity assumes a farmer will not base her input decisions on unexpectedly high or low *future* output, but makes no assumptions about past output. In other words, current and future error terms are unanticipated shocks to productivity. I implement the estimator by taking first-differences to eliminate the fixed-effect and instrumenting the differenced factors with their lagged levels. Lagged levels are uncorrelated with the combined error term by sequential exogeneity, so the instruments are valid.

Table 1.3 reports the Anderson-Hsiao estimates of the production function. As expected, rice farming is relatively labor- and land-intensive. The production function has decreasing returns, and each shock to productivity has the expected sign. The first-stage regressions of factor changes on their lags easily satisfy the usual standards for strength (Stock et al., 2002).

Table 1.3

Production Function Estimates

Production Elasticities		Productivity Modifiers	
Capital		Hunger	-0.032
-Share (θ_K)	0.110		(0.08)
	(0.04)	Illness	-0.147*
-1st Stage F-Stat	111.720		(0.08)
		Death	0.029
Land			(0.12)
-Share (θ_T)	0.244	Flood	-0.101***
	(0.05)		(0.04)
-1st Stage F-Stat	128.387	Pests	-0.057
			(0.04)
Labor		Bad Rain	-0.131***
-Share (θ_L)	0.310		(0.03)
	(0.04)	Low Yield	-0.152***
-1st Stage F-Stat	133.424		(0.02)
		Low Price	-0.065***
Returns to Scale:			(0.02)
-Estimate (σ)	0.664		
	(0.07)		
Households:	734	Observations:	4856
R-Squared:	0.760	Cragg-Donald Stat.	204.435

Note: The table reports the Anderson-Hsiao estimates of the production elasticities and the effects of each component of productivity (see Section 1.5.3 for details). I cluster standard errors by household.

Table 1.4

Sample Sizes and Productivity in Rice-Farming Villages

Farmers Per Village		Productivity Dispersion	
25th Pctl.	5	75/25	1.81
50th Pctl.	9	90/10	3.09
75th Pctl.	12	95/5	3.64

Table 1.5

Correlates of Under-Allocated Farmers

Correlates of Being Under-Allocated in 1996					
Age of Head	Rents Land?	Years Farmed Rice	Illness	High Risk Aversion	Cash Savings
-0.6**	-2.3	-0.5**	-4.6	11.2**	0.0

I define "under-allocated" to mean the household produces more after reallocation than before. I regress a dummy for being under-allocated in the first year (1996) on a set of variables that might cause a household to be allocated too much or too little. The table reports the coefficients and significance levels from the regression. After dropping households for whom all variables are not defined the sample size is 350.

1.5.4 Sample Characteristics

Table 1.4 reports sample sizes and median productivity dispersion among the villages of my sample. The median 90/10 ratio for productivity within a village is 3.09, a number close to the range of 1 to 3 that Gandhi et al. (2011) find for the gross production functions of several manufacturing industries in Colombia and Chile. Productivity in a rice-farming village is distributed much like in a typical manufacturing industry. Hsieh and Klenow (2009) find much larger 90/10 ratios in their sample, most likely because they use value-added rather than gross production functions. Gandhi et al. show that value-added representations require strong assumptions about gross production functions and tend to inflate the dispersion of productivity by an order of magnitude.

1.6 Results

I plug the estimates of production elasticities and anticipated productivity into the expressions for fully perfect and perfect factor market allocations: (1.3) and (1.4) for capital and similar expressions for land and labor. I drop all observations from village-years with only a single farmer because they by construction have no within-village misallocation.[17]

Table 1.5 reports predictors of being under-allocated in the earliest year as per a linear probability regression. I call a farmer under-allocated if he produces more after reallocation than before, and predict under-allocation with the farmer's age and years farming rice,

[17]The sample loses 13 households, 8 villages, and 47 observations.

whether he rents land, whether illness lowered his income, his cash savings, and whether he chose the most risk-averse options in two questions that measure risk preferences.[18] Ten more years of age or experience reduce by 6 and 5 percentage points the chance of being under-allocated, suggesting farmers accumulate factors as they age. As expected, a risk-averse farmer is 11.2 percentage points more likely to be under-allocated. A farmer who fears risk is less likely to gamble on a large farm even if he is more talented than his neighbors. Savings do not predict under-allocation, perhaps because they may equally be a sign of wealth or unwillingness to invest.

I then estimate the misallocation in each village using the expression for G_I in Section 1.3.4 and plot the distribution in Figure 1.2.A. Even in the earliest year of my sample (1996) the total cost of factor and financial market failures was less than 15 percent in most villages. Over time the distribution shifts downward, and misallocation falls below 8 percent for most villages by 2008. Many villages appear to have negative misallocation because estimated misallocation is a random variable. When true misallocation is low the probability a normally distributed estimator falls below zero is high.

To reduce the noise and represent the whole rice sector, Figure 1.2.B depicts misallocation across all villages from the original four provinces surveyed at baseline.[19] I estimate the gains in sample-wide output from reallocating factors within each village after weighting by population (see Section 1.3.5). Sample-wide misallocation is never more than 17 percent and falls to below 4 percent by 2008. The results suggest factor and financial market failures do not produce much costly misallocation, and what little they do produce falls over time.

Figure 1.2.B also separates the misallocation caused by imperfect factor markets from that caused by imperfect financial markets. I calculate G_I^{FACT} and G_I^{FIN} as defined in Section 1.3.5 and reweight them to the sample-level. Both types of misallocation fall from 1996 to 2008. Since the factor market measure is a lower bound while the financial market measure is an upper bound, neither market unambiguously causes more misallocation until 2006 when financial market misallocation drops to nearly zero. Development economists often blame the financial markets for underdevelopment, but the graph suggests factor markets cause as much or even more misallocation. Like in Figure 1.2.A, the apparently negative financial market misallocation in 2008 is an artifact of sampling error. If true financial

[18]Each question asked whether he would accept a gamble that could with equal probability double or reduce to two-thirds his current income. If he refused he was offered a similar gamble where the worse outcome would give 80 percent of his current income. I mark the farmer as highly risk averse if he refused both offers. The question was first asked in the 2003 survey, so I linked the 2003 response to the 1996 status of under-allocation.

[19]I restrict the sample to households from the original four provinces surveyed at baseline to avoid the artificial jump that comes from adding a new province partway through.

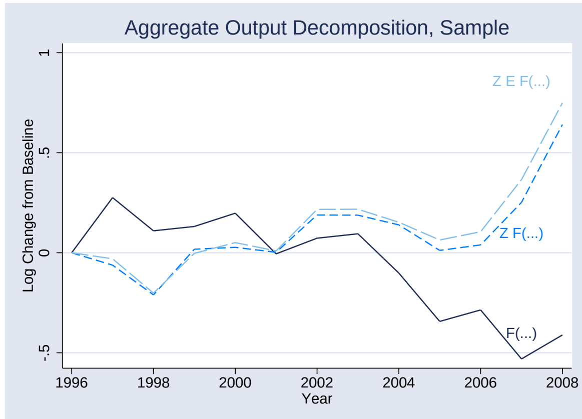
Figure 1.2
Panel A: Density of Within-Village Misallocation; **Panel B:** Sample-Level Overall, Factor, and Financial Market Misallocation

Note: Panel A plots the shifting distribution of misallocation within each village . I report misallocation as the fraction of observed output foregone because factors are misallocated. Panel B calculates the overall cost to the rice sector from misallocation within villages for every year of my sample. It also splits overall misallocation into misallocation from factor versus financial markets, where my measures bound the gains from perfecting first the factor markets and then the financial markets.

market misallocation is almost zero the probability of estimating it to be negative is large. The estimator for financial market misallocation is also more volatile than the estimator for factor market misallocation, so the apparent contrast between the slow march of factor markets versus the drunken stumble of financial markets may be an artifact.

Figure 1.3 decomposes growth in the four provinces into changes in aggregate factor stocks $F(\cdot)$, revenue productivity Z, and the efficiency of factor allocations E. Each line shows how log output would have grown since 1996 if some parts of growth had been shut down. The solid line shows output if productivity and efficiency were fixed at their 1996 level and only aggregate factor stocks changed. Without growth in productivity and efficiency, rice output would have fallen since 1996 as factors flowed out of rice farming. Since Thailand has rapidly industrialized over the past two decades, agriculture's decline is not surprising. The middle dashed line shows growth if changes in productivity are turned back on, and comparing it to the solid line shows the contribution of productivity to growth. Rising productivity since 1998 overwhelmed the outflow of factors and produced net gains in rice revenue. It rose for two reasons: better yields and higher prices. Average yields might have improved as less productive farmers left farming and those who stayed became more skilled, but the spike in productivity after 2006 comes entirely from rising food prices. The final line in Figure 1.3 shows output when changes in efficiency are turned back on, and comparing it to the middle dashed line shows the contribution of improving efficiency

Figure 1.3
Decomposition of Growth in Aggregate Rice Output in the Sample



Note: I decompose aggregate output and compute changes in the log of the aggregate factor stocks $F(\cdot)$, revenue productivity Z, and the efficiency of factor allocations E. Each line plots the counterfactual change in output holding all components except the indicated component fixed (so the lowest line holds average productivity and allocative efficiency fixed while letting aggregate factor stocks change).

to output growth. It is trivial. Compared to the other two sources of growth, efficiency barely changed the trajectory of rice output.

Why do I find so much less misallocation than earlier work (e.g. Hsieh and Klenow, 2009)? One possibility is that markets in agriculture really are more efficient than in manufacturing. The possibility contradicts popular perception that rural villages are backwards, although Townsend (1994) and Benjamin (1992) find surprising efficiency in rural insurance and labor markets. Another possibility is that equally bad markets cause more misallocation in manufacturing than in farming because some feature of manufacturing production makes misallocations more costly. One possibility is that productivity is more dispersed in manufacturing, but I show in Section 1.5.4 that productivity is no less dispersed in a Thai village than a typical manufacturing industry.[20] More generally, the distribution of village-level misallocation in Figure 1.2 shows that some villages have lots of misallocation (nearly 60 percent), demonstrating that costly misallocation is possible in my context; it simply does not happen often. A third possibility is that most misallocation in earlier studies did not come from factor and financial market failures. For example, studies like

[20] Another possibility is that manufacturing has higher returns-to-scale that make misallocation more costly. Consider a simple example: two firms that produce only with capital. Suppose K_1^* and K_2^* are the optimal levels of capital but x units of Firm 2's capital have been inefficiently allocated to Firm 1. Then efficiency is $E = \frac{A_1(K_1^*+x)^\sigma + A_2(K_2^*-x)^\sigma}{A_1(K_1^*)^\sigma + A_2(K_2^*)^\sigma}$ and $\frac{dE}{dx} = -\sigma \frac{\frac{A_2}{(K_2^*-x)^{1-\sigma}} - \frac{A_1}{(K_1^*+x)^{1-\sigma}}}{A_1(K_1^*)^\sigma + A_2(K_2^*)^\sigma} 0$ which is negative because the marginal products are equal at the optimum, and Firm 1 has more capital than is optimal while Firm 2 has less. So efficiency falls faster for higher returns-to-scale, and equivalently misallocation rises faster.

Hsieh and Klenow's import to poor countries parameters calibrated with U.S. data. If rich country production differs from poor country production and the difference grows with the gap in development, then poorer countries might seem to have more misallocation. That may be why Midrigan and Xu (2013) could not explain much misallocation with financial market constraints. Perhaps little of the misallocation Hsieh and Klenow find in India and China comes from factor and financial market failures, an unexpected result given how much attention they receive in the literature on development.

1.7 The Effect of Credit on Misallocation: The Million Baht Program

Between May 2001 and May 2002 the Thai government gave the public lending funds of all the villages in my sample one million baht. The aptly named Million Baht Program in effect gave smaller villages more credit per-household. Kaboski and Townsend (2011) explain that village boundaries have little economic meaning and come from a bureaucratic tangle with statistically random outcomes. Since village sizes are random the per-household rise in credit is also random, and Kaboski and Townsend verify there are no differential trends between the villages that received more or less credit (see Table I on p. 1369 of their paper).

I exploit the program to test my measures of efficiency. By increasing the supply of credit the program improved financial markets and should decrease misallocation from financial market imperfections. [21] I regress a village's misallocation in each year on year dummies, village fixed-effects, the log of the per-household credit injection (one million divided by the number of households), and the interaction between the log credit injection and 2001, the year of implementation, and 2002, the year after. The coefficients on the interactions measure the semi-elasticities of misallocation with respect to credit.

Table 1.6 reports the results, which are rescaled to show the change in the dependent variable due to a one percent increase in per household credit. A one percent increase in credit decreases misallocation by .1 percent of observed output, nearly all of which comes from decreases in financial market misallocation. Since a credit intervention should affect financial markets, the results validate my measures of factor versus financial market misallocation. The program had no significant effect on aggregate land, labor, or capital, confirming Townsend and Kaboski's (2009; 2011) finding that the program did not affect average investment. If the average did not change but misallocation fell, most households

[21] The program might increase misallocation if the village funds lent out credit unfairly. But as Kaboski and Townsend (2011) explain, villagers elected panels of managers to administer the funds. The decisions were transparent and the main criterion was whether the managers thought the borrower could repay the loan.

Table 1.6
Effects of the Million Baht Credit Intervention

	Overall Cost b/se	Cost, Fact. b/se	Cost, Fin. b/se	Agg. K b/se	Agg. T b/se	Agg. L b/se
Log of Credit	0.002	-0.002	0.003	9.766	0.191^{***}	0.542
	(0.05)	(0.01)	(0.04)	(494.52)	(0.06)	(0.72)
2001 X Credit	-0.095^{**}	-0.002	-0.093^{**}	-396.228	0.114	-0.343
	(0.05)	(0.01)	(0.04)	(432.99)	(0.08)	(1.34)
2002 X Credit	-0.028	0.006	-0.034	-221.249	-0.031	-0.147
	(0.04)	(0.01)	(0.04)	(583.20)	(0.09)	(1.58)
Year FEs	X	X	X	X	X	X
Village FEs	X	X	X	X	X	X
Villages	65.000	65.000	65.000	65.000	65.000	65.000
Observations	735.000	735.000	735.000	735.000	735.000	735.000

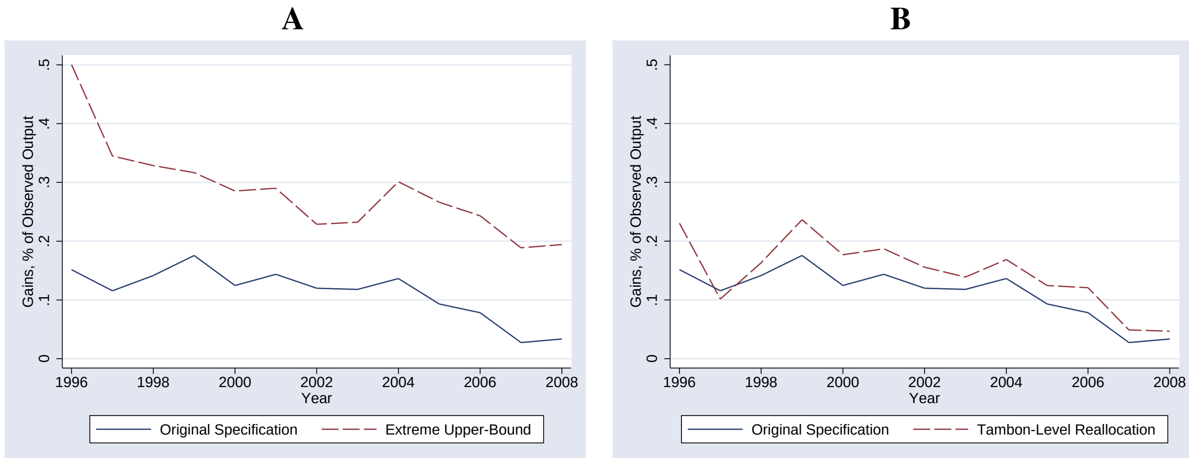
All standard errors clustered by village

Note: The table captures the causal effects of the credit intervention on misallocation. Log of Credit is the log of the per-household credit injection ($\frac{one\ million}{\#\ of\ households}$), and the program was implemented in 2001 and 2002. The coefficient on the interactions gives the percentage point decrease in misallocation caused by a 1 percent increase in per household credit availability. A 1 percent increase in credit reduces overall misallocation by .09 percentage points, and almost all of the effect came from a reduction in misallocation from financial market failures (third column) rather than factor market failures (second column). The program had no effect on (per household) aggregate factor stocks.

must have cut back their scale while the most productive farmers scaled up. Kaboski and Townsend's structural model similarly showed the program did not affect all households equally.

Though statistically significant, the effect is small. Even a fifty percent increase in credit would only reduce misallocation by five percentage points. With so little misallocation at baseline the result is not surprising. Improving financial markets that do not cause much misallocation will not produce spectacular results. But whatever the program's effects, they seem to have faded by its second year. The interaction of average credit injected and the year after implementation (2002) is a third the size and insignificant. One possibility is that households changed how they used their credit in the second year of the program. But recall that measured misallocation is noisy, making the estimated program effect noisy. The impact could be identical in both years, but the variance of the estimators might make it seem different.

Figure 1.4
Panel A: Sample-Wide Misallocation When All Productivity is Anticipated; **Panel B:** Sample-Wide Misallocation with Reallocation Within Sub-districts



Note: Panel A plots misallocation using the breakdown between anticipated and unanticipated productivity in the main text next to misallocation assuming all productivity is anticipated. Panel B plots misallocation assuming reallocation within villages (as in the main text) next to misallocation with reallocation within sub-districts (tambons), most of which contain four villages.

1.8 Robustness

1.8.1 Do Farmers Anticipate More Productivity than Assumed?

I must make assumptions about how much productivity the farmer anticipates to avoid blaming factor and financial markets for random events. If I underestimate how much the farmer knows about his productivity, however, I will underestimate the amount of misallocation. I can bound the bias by recalculating misallocation under an extreme assumption: farmers anticipate all productivity. The allocation I calculate by subbing $A_{it}\phi_{it}$ in for A_{it} in expression (1.3) will produce an unrealistically perfect outcome because it assumes farmers know everything: not just their own managerial talent, but also how much of the harvest rats will eat in three months. Given even the extreme upper-bound on misallocation, do I find as much misallocation as Hsieh and Klenow (2009)?

Figure 1.4.A, which compares the upper bound to my preferred specification, suggests otherwise. Misallocation is higher in all years but still falls short in most years of the 40 to 60 percent Hsieh and Klenow report India could gain from reallocating to U.S. efficiency, much less the 100 percent it could get with perfect efficiency. By the end of the sample misallocation costs less than 20 percent of observed output—half of what Hsieh and Klenow find for the U.S.

Table 1.7
Correlation Between Village Misallocation and Sample Size

	(1) Efficiency b/se	(2) Gains b/se
Number of Farmers	0.000	-0.004*
	(0.00)	(0.00)
Villages	65	65
Observations	735	735

Note: I regress efficiency and gains from reallocation with the village on the number of farmers in the village in my sample. I exclude villages with only one farmer (they have zero misallocation by construction). I find no evidence that small samples bias me towards finding too little misallocation. I cluster standard errors at the village-level.

1.8.2 Does a Small Sample Cause Underestimates of Misallocation?

Misallocation often happens because the most talented producers do not get enough factors. If very talented farmers are rare they might not appear in my small per-village samples and I might underestimate misallocation. Even if the social planner favors reallocating everything to a small productive elite, I show in Section 1.5.4 that the dispersion of productivity in most villages is similar to what Gandhi et al. (2011) find in manufacturing industries. Table 1.7 reports regressions of efficiency on sample size. Estimated efficiency is no lower in villages with larger samples, and equivalently the gains from reallocation are no higher in villages with larger samples (if anything, they are lower). The results give no reason to fear I would find more misallocation if my sample were larger. In Figure 1.4.B I recalculate sample-wide misallocation assuming I can reallocate factors between villages within a sub-district. The procedure does not differ much from within-village reallocation except I must account for the sample design (see Appendix A.2). Since sub-districts contain several villages they have larger samples of farmers. Reallocating within sub-district also increases the potential gains because the social planner can reallocate between as well as within villages. Figure 1.4.B shows the combined effect of both forces is small. Reallocating between villages does not much increase misallocation. Sample size is not a problem, and between-village misallocation is small relative even to within-village misallocation.

1.9 Conclusion

Imperfect factor and financial markets do not cause much misallocation among Thai rice farmers. Of the little they do cause, factor markets cause as much or more than financial markets. Declining misallocation makes positive but trivial contributions to growth in aggregate rice output, whereas most growth is caused by the overall flow of factors out of rice production and an overall increase in average farm productivity. The Thai government's Million Baht credit program did raise efficiency through the expected channel: a statistically significant decrease in misallocation from poor financial markets. But the effects were small, most likely because there was so little misallocation to begin with.

The results do not mean misallocation never matters or market failures have no costs. The assumptions behind my method might hold for other crops and other countries, but that does not mean applying it to India or Africa would reveal as little misallocation as I find in Thailand. It is also possible, as I discussed in Section 1.6, that factor and financial market failures cause more misallocation outside agriculture. Finding little misallocation within rice farming is not the same as finding little misallocation between farming and industry, and Jeong and Townsend (2007),Buera et al. (2011), and Midrigan and Xu (2013) seem to find between-sector misallocations do matter. Even among Thai farmers, I find in Shenoy (2014b) that risk causes costly under-specialization in economic activities.

What the results do suggest is that evidence of a market failure is not always evidence of costly misallocation. Only by measuring the costs of each type of misallocation caused by each market failure can we know when and where misallocation really matters.

CHAPTER 2

Risky Income or Lumpy Investments? Testing Two Theories of Under-Specialization

2.1 Introduction

> *To take...the trade of the pin-maker...One man draws out the wire, another straights it, a third cuts it...ten persons, therefore, could make among them upwards of forty-eight thousand pins in a day...But if they had all wrought separately and independently...they certainly could not each of them have made twenty, perhaps not one pin in a day...*
>
> -Adam Smith, *Wealth of Nations*

The idea that specialization is efficient is as old as economics itself. The puzzle, then, is to explain why households in poor countries rarely specialize in a single business or a single job (Banerjee and Duflo, 2007). If entering multiple economic activities is costly, why would the world's poorest people fail to specialize?

This paper tests two well-known but unproven theories for why the poor have so many economic activities: the theory of risky income and the theory of lumpy investments. The theory of risky income compares a poor household choosing economic activities to an investor choosing stocks. Like stocks the activities of the poor have risky returns, driving households to diversify their portfolio even though expanding it is costly. Whereas this theory blames under-specialization on a lack of insurance, the theory of lumpy investments blames a lack of credit. The theory posits that households must make a large investment—tailors must buy a sewing machine and bakers must buy an oven—before expanding any business to its optimal scale. Households that cannot borrow enough to create one large business must cobble together income from many small businesses.

From a simple model I derive four tests of the theory of risky income. Each household has a primary activity and pays a fixed cost to enter any side activity. The returns to these activities are random and not perfectly correlated. Therefore the theory's first test is that a rise in the riskiness of the primary activity causes a risk-averse household to self-insure by entering more side activities. But since labor spent on side activities is labor taken from the primary activity, a rise in the average return to the primary activity raises the cost of self-insurance. The theory's second test is that a rise in the return to the primary activity causes the household to exit side activities.

The third test, which uses revenue from side activities rather than total revenue to check whether specialization is efficient, is critical for two reasons. First, my empirical approach rules out any test using total revenue. Second, the household reallocates labor between activities when it enters new activities. For both reasons I cannot identify the fixed cost of entering a new activity, the clearest sign of inefficient under-specialization. I can, however, derive the optimal allocation of labor as a function of the number of activities, and use this allocation to find the change in side revenue caused when the household enters a new activity. I show that if the change is negative then specialization is inefficient, though the converse need not hold.

To run these tests I study how rice farmers in Thailand respond to volatility in the international price of rice. Using a monthly panel I identify the households who expect a rice harvest in the next three months. Higher volatility in the price of rice raises the riskiness of these farmers' income. By comparing their response to the response of farmers who do not expect a harvest I identify the causal effect of riskier income on specialization. By likewise comparing how the two groups respond to changes in the expected price I identify the causal effect of greater returns on specialization. My first two tests confirm the theory of risky income. Greater risk drives households into more activities while higher returns tempt households out of activities. After adjusting for how well international prices predict local prices, my baseline estimates suggest a 21 percent rise in volatility causes a household to enter an extra activity.

Since a household that expects a harvest next month sells no rice this month the mean and variance of the rice price change the number of activities without directly affecting current revenue. I use this change to instrument for the number of activities. Households expecting a harvest do not yet have the revenue from their primary activity, ruling out any test of whether additional activities decrease total revenue. But my third test shows that if additional activities decrease revenue from side activities then a failure to specialize is inefficient. Two-stage least squares confirms exactly that.

Finally, I test the theory of lumpy investments. The theory predicts that households with easier access to credit can afford the lumpy investments that let them specialize in one business. I test whether households exit activities after a government program creates random variation in the availability of credit, but find no evidence that credit increases specialization.

Existing work links risk to under-specialization but lacks the exogenous variation needed to show that risk causes under-specialization.[1] Morduch (1990) shows that households more vulnerable to income shocks tend to diversify their crops, and Bandyopadhyay and Skoufias (2012) find that households in areas with riskier weather tend to have spouses with different occupations. But since vulnerability and weather risk are not exogenous, households who endure these problems may endure other problems unrelated to the riskiness of their income. If these other problems also cause under-specialization then estimates of the effect of risk on specialization will be biased.

More recent work, on the other hand, uses exogenous variation but does not study the effect of risk on specialization. Karlan et al. (2012) find that randomly assigned weather insurance increases both investment and the proportion of land planted with crops sensitive to rain, but they do not study specialization. Adhvaryu et al. (2013) study whether households expand their number of activities in response to shocks, but entering activities in response to shocks is not the same as entering activities in anticipation of risk.

Unlike the existing literature, this paper uses plausibly exogenous variation to identify whether risk causes costly under-specialization. I identify the effect of ex ante risk rather than ex post shocks. Finally, to my knowledge this paper is the first to test the theory of lumpy investment alongside the theory of risky income, letting me assess which theory has more merit.

2.2 Theory: A Model of Risky Income

2.2.1 Deriving Tests for the Theory

Each household has one primary economic activity and may enter any number of side activities. The household pays a fixed cost for each side activity. It allocates one unit of labor between all activities. Labor produces a constant return, and the household does not know the return to any activity until after it has made its choices.

[1] Many more papers study how imperfect insurance drives households to make other inefficient choices. Those most relevant to this paper whether farmers with riskier profits marry their daughters to men in different occupations (Rosenzweig and Stark, 1989), choose safer but less profitable bundles of investments (Rosenzweig and Binswanger, 1993; Bliss and Stern, 1982), or delay the planting of their crops (Walker and Ryan, 1990, p. 256).

The household must first choose the number of side activities. Then it chooses the allocations of labor. Then the returns to the side activities are realized. Finally, the household learns the return to its primary activity and consumes.

Suppose for simplicity that the household has constant absolute risk-averse preferences. The household solves

$$\max_{M,L_p,\{L_{s,m}\}} \mathbb{E}[-e^{-\alpha C}]$$

subject to

$$C = Y = w_p L_p + \sum_{m \in M} w_{s,m} L_{s,m} - MF$$

$$L_p + \sum_m L_{s,m} = 1$$

where α is the coefficient of absolute risk aversion, $M \geq 0$ is the number of side activities, and L_p and $\{L_{s,m}\}_{m \in M}$ are the labor allocated to the primary and each side activity. The household consumes its revenue, which is the sum of revenue from primary (p) and side (s) activities minus fixed costs. The primary and side activities yield returns w_p and $\{w_{s,m}\}_{m \in M}$, which are independent normal random variables with $w_p \sim N(\bar{w}_p, \sigma_p^2)$ and $w_{s,m} \sim N(\bar{w}_s, \sigma_s^2)$ for each m.[2] Assume the side activities yield weakly lower expected returns: $\bar{w}^p \geq \bar{w}^s$. Also assume the average premium to the primary activity, $\bar{w}_+ = \bar{w}_p - \bar{w}_s$, is not too large: $\bar{w}_+ < \alpha\sigma_p^2$. If this assumption fails the household will specialize despite the risk.

I make many simplifying assumptions about functional forms, but the important results rest on four crucial assumptions. First, the household is risk-averse. Second, the household cannot perfectly smooth its consumption through insurance or savings. (To sharpen the model's predictions I assume the household has no insurance or savings.) Third, the returns to side activities are not perfectly correlated with returns to the primary activity. Fourth, each activity has (locally) increasing returns. The first two assumptions force the household to insure itself against risk without using financial markets. The third assumption makes under-specialization a form of insurance. The fourth assumption makes under-specialization costly.

[2] If the returns to side and primary activities were not independent, the properties of normal random variables let me write the returns to each side activity as $w_{s,m} = \rho^m w_p + \xi^m$ for some correlation coefficient ρ^m. If I then re-label variables accordingly, all the results should go through. I only need the returns to be imperfectly correlated.

To get the intuition of the model, consider the simple case where the household either specializes ($M = 0$) or has one side activity ($M = 1$). The household chooses between two "bundles" of average consumption $\bar{C}$ and variance of consumption V:

	$M = 0$	$M = 1$
$\bar{C}$	$\bar{w}_p$	$\bar{w}_p - \bar{w}_+(1 - L_p) - F$
V	σ_p^2	$(L_p)^2\sigma_p^2 + (1 - L_p)^2\sigma_s^2$

Since $L^p < 1, \bar{w}_+ > 0$ and $F > 0$ the household can lower the variance of its consumption by entering a side activity if it accepts a lower expected consumption.

Suppose the household enters a side activity and must now choose how much labor to shift from the primary activity. Since consumption is a normal random variable, expected utility is (the negative of) a log normal random variable. The household now solves

$$\max_{L^p} -e^{-\alpha\bar{C} + \frac{\alpha^2}{2}V}.$$

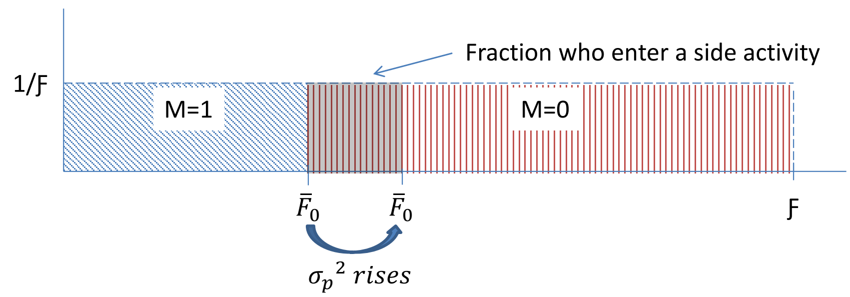
The first-order condition is

$$
\begin{aligned}
0 &= -e^{-\alpha\bar{C} + \frac{\alpha^2}{2}V} \cdot \left(-\alpha\frac{\partial\bar{C}}{\partial L_p} + \frac{\alpha^2}{2}\frac{\partial V}{\partial L_p}\right) \\
\Rightarrow \qquad 0 &= -\bar{w}_+ + \alpha L_p\sigma_p^2 - \alpha(1 - L_p)\sigma_s^2 \\
\Rightarrow \qquad L_p &= \frac{\alpha\sigma_s^2 + \bar{w}_+}{\alpha\left(\sigma_p^2 + \sigma_s^2\right)}
\end{aligned}
$$

To derive predictions about aggregate statistics, suppose the fixed cost of entering the side activity varies across households because some find it easier to enter activities. For example, two rice farmers might differ only in how closely they live to a construction site where they can find part-time work. For simplicity suppose $F \sim U[0, \mathcal{F}]$ for some upper-bound $\mathcal{F}$.

For any amount of risk there is a household whose fixed cost makes it indifferent between zero and one side activity. Call that household's fixed-cost $\bar{F}_0$. Let $C(M)$ and $V(M)$ be the mean and variance of consumption as functions of the number of side activities. Then $\bar{F}_0$ is defined as the fixed cost that makes this equation hold:

Figure 2.1
Intuition of the Simplified Case



Note: M is the number of side activities; $\bar{F}_0$ the threshold fixed cost for moving from zero to one side activity; σ_p^2 is the variance of the primary economic activity. A rise in the variance raises the threshold fixed cost, which is how much households will to pay for insurance, so the highlighted mass of individuals switches from specialization to having a side activity.

$$
\begin{aligned}
& & -e^{-\alpha\bar{C}(0)+\frac{\alpha^2}{2}V(0)} = -e^{-\alpha\bar{C}(1)+\frac{\alpha^2}{2}V(1)} \\
&\Rightarrow & -\alpha\bar{C}(0) + \frac{\alpha^2}{2}V(0) = -\alpha\bar{C}(1) + \frac{\alpha^2}{2}V(1) \\
&\Rightarrow & \frac{\alpha}{2}[V(0) - V(1)] = \bar{C}(0) - \bar{C}(1)
\end{aligned}
$$

Substitute the expressions from the table above and from the optimal labor allocation:

$$\bar{F}_0 = \frac{\left(\alpha\sigma_p^2 - \bar{w}_+\right)^2}{2\alpha\left(\sigma_p^2 + \sigma_s^2\right)}$$

Households who pay fixed costs above the threshold $\bar{F}_0$ will specialize while those below enter a side activity. The threshold rises with the variance of the primary activity σ_p^2, and Figure 2.1 shows the effect on the number of households with a side activity. When their primary activity becomes riskier, households are willing to pay a bigger fixed cost to make their revenue less risky. The threshold $\bar{F}_0$ rises, and the mass of households with fixed costs between the old and new thresholds enter a side activity. The change in the average number of activities in the sample is

$$\frac{\partial\mathbb{E}_F[M]}{\partial\sigma_p^2} = \frac{\sigma_p\left(\alpha\sigma_p^2 - \bar{w}_+\right)\left(\alpha\sigma_s^2\left(\sigma_p^2 + 2\sigma_s^2\right) + \bar{w}_+\sigma_s^2\right)}{\alpha\sigma_s^2\left(\sigma_p^2 + \sigma_s^2\right)^2} \cdot \frac{1}{\mathcal{F}} > 0.$$

(Recall that by assumption $\alpha\sigma_p^2 - w_+ > 0$). Similarly we can derive the change in the average number of activities when the average return to the primary activity rises. Since a rise in the expected return makes under-specialization more costly, the threshold will fall and the average number of activities will fall:

$$\frac{\partial \mathbb{E}_F[M]}{\partial \bar{w}_p} = -\frac{\alpha\sigma_p^2 - \bar{w}_+}{\alpha\left(\sigma_p^2 + \sigma_s^2\right)} \cdot \frac{1}{\mathcal{F}} < 0$$

The intuition of the case where $M \in \{0, 1\}$ holds for any number of activities $M \in \{0, 1, 2, \ldots\}$, and the simple but tedious proof is left for Appendix [X].

Taking on additional side activities lowers the household's expected total revenue. But the empirical approach of Section 2.3 studies rice farmers who expect but have not yet collected a harvest. These farmers do not have their total revenue, ruling out any test based on total revenue. I must instead derive the model's predictions about what under-specialization does to revenue from side activities; that is, what happens to the rice farmer's revenue from cassava when he starts baking bread.

Consider the household's revenue just before it gets the output from its primary activity. Its revenue at this stage is simply the revenue from its side activities:

$$y_s = \sum_{m \in M} w_{s,m} L_{s,m} - MF$$

For simplicity treat the number of activities M as continuous. Holding a household's cost of additional activities fixed, a small increase in the number of activities changes side revenue on average by

$$\begin{aligned}
\mathbb{E}_F\left[\frac{\partial y_s}{\partial M}\right] &= \mathbb{E}_F\left[\mathbb{E}_{w_p,\{w_{s,m}\}}\left[\frac{\partial y_s}{\partial M} \mid F\right]\right] \\
&= \mathbb{E}_F\left[\frac{\partial}{\partial M}\left[-MF + (1 - L_p)\bar{w}_s\right]\right] \\
&= -\mathbb{E}[F] + -\bar{w}_s \frac{\partial L_p}{\partial M}
\end{aligned}$$

The average change in side revenue, which corresponds to the instrumental variables coefficient estimated in Section 2.5.1, has two parts: the average fixed cost of a side activity and the effect on side revenue of shifting labor to the side activities. Since an all-else-equal increase in the number of activities makes the portfolio of side activities less risky, the household wants to shift labor away from its primary activity. Then $\frac{\partial L_p}{\partial M} < 0$ and the second term is positive. If large enough it will swamp the cost of under-specialization and make the derivative (and thus the instrumental variables estimate) positive. To see why, suppose the household starts with no side activities and thus no revenue from side activities. If the variance of the primary activity rises sharply and the cost of entering a side activity is small, then the household will want to enter the side activity. Then revenue from side activities will have increased, and though the increase might be small compared to what the household loses from its primary activity, the coefficient I estimate will be positive. Thus a negative estimate is sufficient evidence that activities (and thus under-specialization) is costly, but it is not necessary evidence. This arguement ignores the direct effect that my instruments, the variance and the average returns, have on the labor allocation. But as I show in the proof in Appendix B.1, the direct effect only reinforces the result.

The model also makes a prediction about the ordinary least squares coefficient, which estimates the average effect of increasing the number of activities without holding their cost fixed. That is, it estimates the average total derivative

$$\begin{aligned}\mathbb{E}\left[\frac{dy^s(M,F)}{dM}\right] &= \mathbb{E}\left[\frac{\partial y^s}{\partial M} + \frac{\partial y^s}{\partial F}\cdot\frac{\partial F}{\partial M}\right] \\ &= \mathbb{E}\left[\frac{\partial y^s}{\partial M}\right] + \mathbb{E}\left[\frac{\partial y^s}{\partial F}\cdot\mathbb{E}\left[\frac{\partial F}{\partial M}\mid M\right]\right] \\ &= \mathbb{E}\left[\frac{\partial y^s}{\partial M}\right] + \mathbb{E}\left[\frac{\partial y^s}{\partial F}\cdot\frac{\partial}{\partial M}\mathbb{E}\left[F\mid M\right]\right]\end{aligned}$$

The term $\frac{\partial y^s}{\partial F}$ is clearly negative; a higher fixed cost will lower revenue. The term $\frac{\partial}{\partial M}\mathbb{E}\left[F\mid M\right]$ gives the selection bias. It captures the difference in fixed cost paid by households who select into many versus few activities. As Figure 2.2 illustrates, it is also negative. Since a household takes up a large number of activities if it pays a small fixed cost, the number of activities is informative about their cost.[3] This gives the final test of the model:

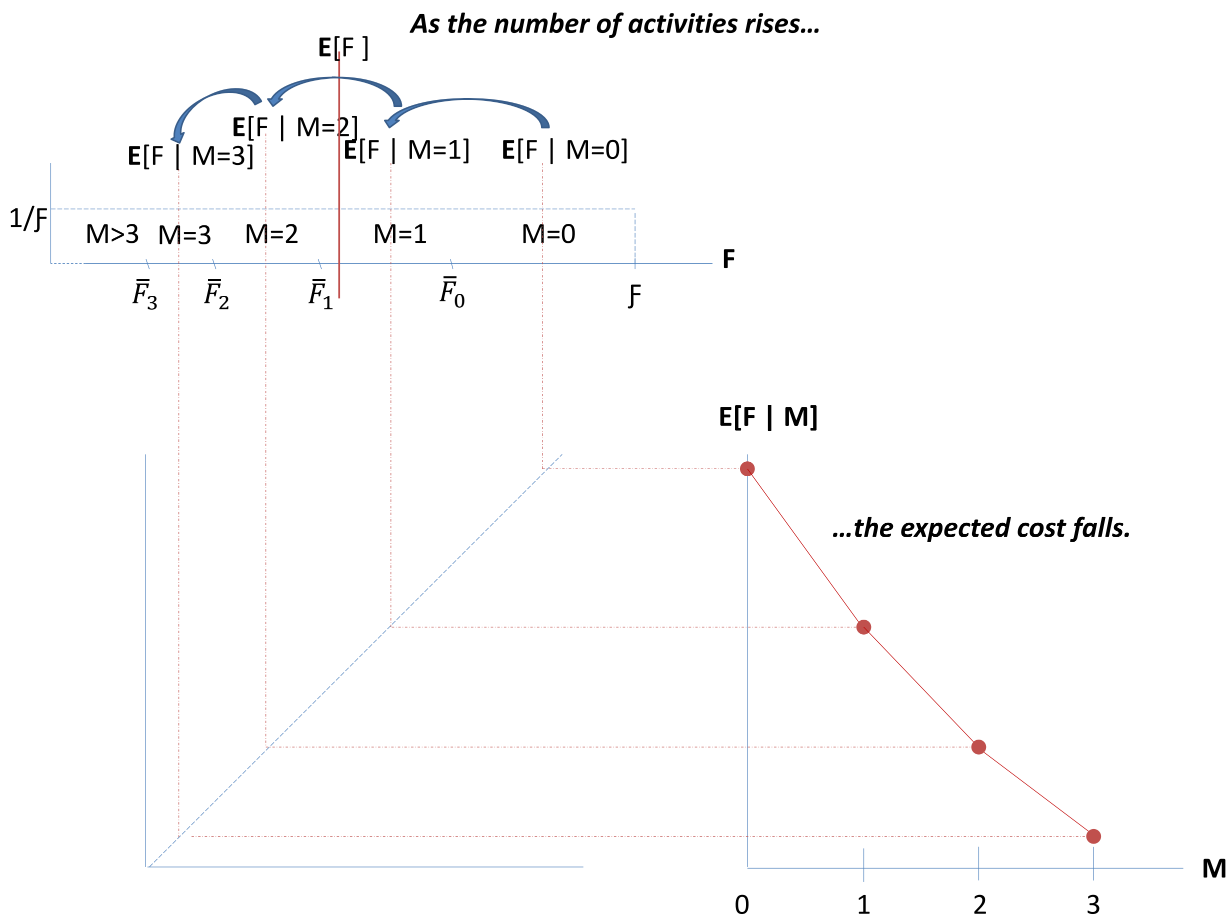
$$\begin{aligned}\beta_{OLS} &= \mathbb{E}\left[\frac{\partial y^s}{\partial M}\right] + \mathbb{E}\left[\frac{\partial y^s}{\partial F}\cdot\frac{\partial}{\partial M}\mathbb{E}\left[F\mid M\right]\right] \\ &> \mathbb{E}\left[\frac{\partial y^s}{\partial M}\right] \\ &= \beta_{IV}\end{aligned}$$

To summarize, the model gives four tests for the theory of risky income:

Test 1 (Risk) *Households enter activities when the returns to their primary activity get riskier.*

[3]Indeed, $\mathbb{E}\left[F\mid M\right]$ is just the demand curve for insurance through under-specialization. Like any demand curve its slope is negative.

Figure 2.2
Why is OLS Upward-Biased?



Note: M is the number of side activities; $\bar{F}_m$ is the threshold fixed cost below which a household moves from m to $m+1$ side activities. A household only enters many activities if its fixed cost low, so the number of activities predicts a household's costs. The costs of extra activities appear in the error term of a regression of side revenue on number of activities, so the coefficient on number of activities is biased upwards.

Test 2 (Return) *Households exit activities when the (expected) returns to their primary activity rise.*

Test 3 (Cost) *The average effect of more activities on revenue is negative only if under-specialization is costly.*

Test 4 (OLS Bias) *Compared to the IV estimate, the OLS estimate of the effect of more activities on side revenue is biased positively.*

2.2.2 Modeling and Measuring Expectations about Risk

To run these tests I must model rice farmers' expectations about the returns and volatility of the price of rice. Suppose the household makes its choices at the beginning of period t. It has not yet observed the price w_{pt} and must form its expectation $\bar{w}_{pt}$ using only information from the past. Suppose the monthly price follows the Autoregressive Conditional Heteroskedasticity (ARCH) model of Engle (1982) with one modification: I assume the level of the price follows a random walk. The assumption reduces the number of parameters I must estimate and, as I show below, matches the true series well. Then

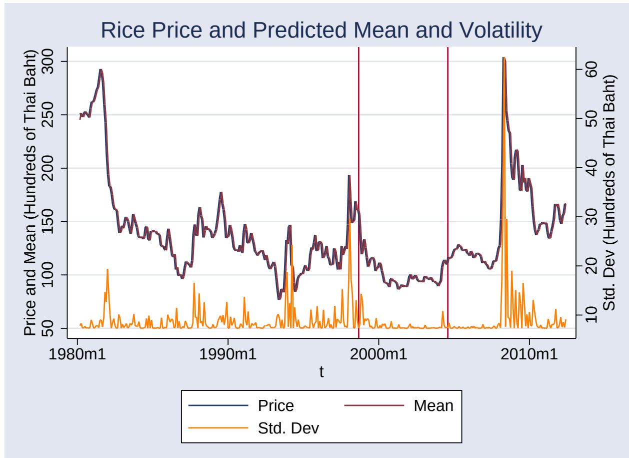
$$\begin{aligned} w_{pt} &= w_{p,t-1} + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t}, \qquad z_t \sim N(0,1) \\ h_t &= \tau_0 + \tau_1 \varepsilon_{t-1}^2. \end{aligned}$$

At the beginning of period t, the household expects a return of $\bar{w}_{pt} = \mathbb{E}[w_{pt}] = w_{p,t-1}$. The variance of the return is $\sigma_{pt}^2 = V(w_{pt}) = V(\varepsilon_t) = h_t = \tau_0 + \tau_1 \varepsilon_{t-1}$. I estimate the model using conditional maximum likelihood.[4] The predicted value $\hat{h}$ is a consistent estimate of the true conditional variance.

In practice I must make several simplifications when I use this measure. I cannot use the actual expected volatility of the price at harvest because the empirical design in Section 2.3 compares farmers expecting a harvest to non-farmers and farmers who do not expect a harvest. Since I cannot define the volatility at the time of harvest for non-farmers I must use the current volatility. This creates measurement error and may bias my estimates towards zero. I also measure volatility using the conditional standard deviation $\sqrt{h}$ rather than the conditional variance to make the coefficients on the volatility of the price and the expected price comparable.

[4]The true distribution of z_t need not be normal; the (quasi) maximum likelihood estimator based on a normal distribution is still consistent.

Figure 2.3
Rice Price and Predicted Mean and Conditional Standard Deviation



Note: I plot the actual rice price together with the predicted rice price and the predicted volatility (square root of the predicted conditional variance) from the Autoregressive Conditional Heteroskedasticity (ARCH) model. The red lines mark the start and end of the panel data.

Figure 2.3 plots the actual price of rice, the predicted mean, and the predicted standard deviation. Simple though it is, the random walk assumption makes very accurate predictions about the mean. A regression of price on its lag gives a coefficient of .995. The red lines mark the start and end of the time period covered in the monthly panel data. The sample spans a time when prices are relatively stable, ending well before the massive food price spike of 2008. [5]

2.3 Empirical Design: Implementing the Tests

2.3.1 Estimating Risk Response

Changes in the international price of rice—and the responses they evoke in Thai rice farmers—provide the exogenous variation in risk I need to test the model. Between planting and harvest the price can change drastically, and anecdotal evidence suggests farmers follow it closely in newspapers, radio broadcasts, and television reports. Since most of my sample grows at least some of the white rice and jasmine rice that make Thailand the

[5]The econometrically-minded may worry if regressions on a regressor generated from a time series model are consistent. Pagan (1984) confirms that the ARCH predicted value (though not the residual) will give consistent estimates, and I have confirmed in monte carlo simulations that panel estimators are consistent as well.

Table 2.1
Rice Prices and Sales

	(1) Avg. Transaction Price	(2) Rice Sold
Int. Rice Price	0.333**	
	(0.14)	
Rice Harvested		0.856***
		(0.01)
Constant	1.500	-2043.744***
	(1.53)	(70.44)
N	62	2126

Note: **Column 1 —** The dependent variable is the sample-wide average price of a kilogram of rice based on actual transactions, and the independent variable is the international price of rice in baht per kilogram. Not all survey rounds include any sales of rice—hence the number of observations is smaller than the number of survey rounds. Standard errors are robust to heteroskedasticity. **Column 2 —** The unit of observation is the household-month conditional on positive rice harvest.

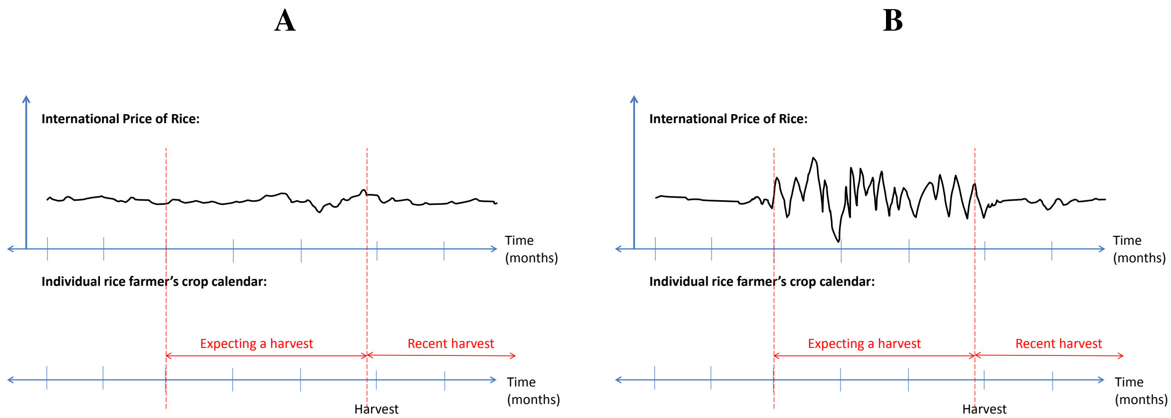
world's biggest rice exporter, the international price matters.[6] In Column 1 of Table 2.1 I report the correlation between the sample-wide average price farmers receive and the international price. Though not perfect, the correlation is significant and large enough to make following the international price worth a farmer's time. If prices become more volatile the farmers know it and know the value of their harvest has become riskier.

A response to volatility need not be a response to risky income unless it comes from a specific group of farmers: those who harvest soon. Simply comparing the response of a rice farmer to someone who does not farm rice might just measure how rice farmers differ in their attitude to risk. Observing a household with rice planted but not yet harvested—a farmer expecting a harvest in the next three months —isolates the effect of risky income. Farmers harvest rice roughly four months after planting and cannot hasten or delay the date. Harvesting too soon yields immature grains while harvesting too late risks losses to pests. The International Rice Research Institute states that "the ideal harvest time lies between 130 and 136 days after sowing for late" varieties and gives similarly narrow windows for other varieties (Gummert and Rickman, 2011). Leaving rice on the stalk to wait out low prices is not an option.

Although in principle a farmer might store rice after harvesting, threshing, and drying, in practice the farmers in my sample sell most of their rice as soon as they harvest it. Colum 2 of Table 2.1 reports the correlation between how much rice a household sells

[6]As expected, I find in unreported regressions that farmers harvesting only sticky rice have a lower response. The negative response too large to be the all-else-equal effect of growing rice that will not be exported. Households who grow only sticky rice are unusual, their response may differ from that of other farmers for reasons beyond the type of rice.

Figure 2.4
Response to Conditional Volatility



Note: Among rice expecting a harvest, I compare the response when rice prices are (**A**) stable to when they are (**B**) volatile. Since I use household fixed-effects I effectively compare each farmer to himself.

and how much it harvests conditional on harvesting any during the month. It suggests farmers sell almost every kilogram of rice as soon as it comes from their fields. Households either cannot arbitrage—perhaps because millers and other middlemen only buy at certain times—or they need cash too desperately to wait.

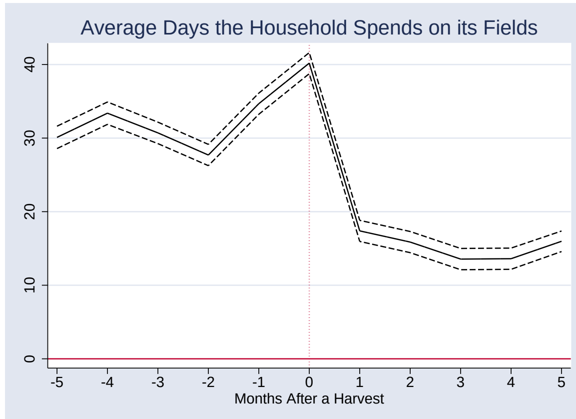
The farmers in my sample are too small to affect the international price and they cannot delay their harvest. After controlling for the responses of non-rice farmers and rice farmers not expecting a harvest, any additional response a farmer makes to higher volatility just before her harvest must be caused by riskier income. Since I have a panel I can also control for household fixed-effects to eliminate any fixed source of bias.[7] The regression I run will actually compare the farmer to herself at times when prices are volatile but she expects no harvest, and times when she expects a harvest but prices are not volatile. Figure 2.4 illustrates the specification.

When prices become volatile the farmer must decide whether to shift her efforts away from maximizing the upcoming harvest. Figure 2.5 graphs the average household labor that rice farmers devote to their fields in the months before and after harvest. Bringing a rice crop to harvest requires constant effort right up through harvest. Working as a laborer or planting cassava detracts from rice farming just like that model assumes extra side activities detract from the primary activity.

To estimate the response I run the regression

[7] For example, suppose only rich farmers plant in January to harvest in April. If the rice price always turns volatile in March, I might just estimate the effect of being a rich farmer. If seasonal selection and seasonal volatility matter, fixed-effects will deal with them.

Figure 2.5
An Impending Rice Harvest Requires Labor



Note: The figure shows how many days the average household works in its fields in the months before and after a rice harvest. More precisely, I plot the coefficients of a regression of the number of days worked in the fields on dummies for periods before and after the harvest. The dashed lines cover 95 percent confidence intervals. A rice crop requires labor in the months leading to the harvest.

$$\begin{aligned}[Activities]_{it} = {} & [FE]_i + \beta_M[Mean]_t + \beta_V[Volatility]_t \\ & + \beta_E[Expecting\ Harvest]_{it} + \beta_H[Had\ Harvest]_{it} \\ & + \beta_{RM}[Rice\ Farmer]_i \times [Mean]_t + \beta_{RV}[Rice\ Farmer]_i \times [Volatility]_t \\ & + \beta_{EM}[Expecting\ Harvest]_{it} \times [Mean]_t + \beta_{EV}[Expecting\ Harvest]_{it} \times [Volatility]_t \\ & + \beta_{HM}[Had\ Harvest]_{it} \times [Mean]_t + \beta_{HV}[Had\ Harvest]_{it} \times [Volatility]_t + \varepsilon_{it}.\end{aligned} \tag{2.1}$$

Aside from the responses of non-farmers and farmers who do not expact a harvest, I must also control for the responses of farmers who just had a harvest. Having had a harvest is negatively correlated with expecting a harvest and cannot be left in the error term. In some regressions I replace the main effects [Mean] and [Volatility] with month dummies. Month dummies eliminate much of the variation in volatility but produce more conservative estimates. Since the volatility is generated I use a two-stage bootstrap for all inference in the results I report in Section 2.5.1. The details of the bootstrap are in the Online Appendix B.3.

The coefficient β_{EV} on $[ExpectingHarvest] \times [Volatility]$ measures the average response to volatility of a farmer who expects a harvest, while controlling for the responses of non-farmers and farmers without upcoming harvests. Since the number of activities is my measure of specialization, β_{EV} measures the causal effect of risk on under-specialization. Test 1 predicts it should be positive. The coefficient β_{EM} on $[ExpectingHarvest] \times [Mean]$ measures the response to higher average prices, and Test 2 predicts it should be negative.

2.3.2 The Costs of Under-Specialization

Risk may drive households into extra side activities, but are they costly? It is hard to imagine why else the household would diversify only when risk increases. If the extra activities were costless the household ought to have as many as possible. Test 3, however, suggests a direct approach: to check whether revenue from side activities falls as the farmer adds more activities.

Rises in volatility will cause farmers expecting a harvest to increase their number of activities, but by construction these farmers have not yet sold their harvest and collected their primary revenue. I cannot run any test on total revenue. Test 3 solves the problem by showing that if revenue from side activities falls when the household adds activities then under-specialization is costly. Figure 2.6 shows that most rice farmers have many activities—cassava fields or bakeries—even when volatility is average. Test 3 says that if a rice farmer's revenue from cassava falls when he baking bread, and the loss to cassava outweighs the gain from bread, then extra activities are costly. I can confirm extra activities are costly if I have instruments that drive farmers into more activities without directly affecting revenue.

The response of farmers expecting a harvest to the mean and volatility of prices is exactly the instrument I need. Since household revenue before the harvest does not include revenue from rice, movements in the rice price cannot affect revenue directly. Greater risk might cause a household to invest less in physical and human capital, but the effect will not appear for years to come. My regressions measure very short run changes from month-to-month. I can then run the following first-stage regression

$$
\begin{aligned}
[Activities]_{it} = [FE]_i + \sum_m \beta_{D,m}[Month\ Dummy] \qquad (2.2) \\
+ \beta_E[Expecting\ Harvest]_{it} + \beta_H[Had\ Harvest]_{it} \\
+ \beta_{RM}[Rice\ Farmer]_i \times [Mean]_t + \beta_{RV}[Rice\ Farmer]_i \times [Volatility]_t \\
+ \beta_{EM}[Expecting\ Harvest]_{it} \times [Mean]_t + \beta_{EV}[Expecting\ Harvest]_{it} \times [Volatility]_t \\
+ \beta_{HM}[Had\ Harvest]_{it} \times [Mean]_t + \beta_{HV}[Had\ Harvest]_{it} \times [Volatility]_t + \varepsilon_{it}.
\end{aligned}
$$

The second-stage regression excludes $[ExpectingHarvest] \times [Mean]$ and $[ExpectingHarvest] \times [Volatility]$ like so:

$$
\begin{aligned}
[Revenue]_{it} = [FE]_i + \gamma_A[\widehat{Activities}]_{it} + \sum_t \gamma_{D,m}[Month\ Dummy]_t \qquad (2.3) \\
+ \gamma_E[Expecting\ Harvest]_{it} + \gamma_H[Had\ Harvest]_{it} \\
+ \gamma_{RM}[Rice\ Farmer]_i \times [Mean]_t + \gamma_{RV}[Rice\ Farmer]_i \times [Volatility]_t \\
+ \gamma_{HM}[Had\ Harvest]_{it} \times [Mean]_t + \gamma_{HV}[Had\ Harvest]_{it} \times [Volatility]_t + u_{it}.
\end{aligned}
$$

Test 3 predicts the coefficient on $[Activities]$ γ_A should be negative. The final test, Test 4, predicts the coefficient on $[Activities]$ κ_A in the simple OLS regression

$$
[Revenue]_{it} = \kappa_A[Activities]_{it} + \sum_t \kappa_{D,m}[Month\ Dummy]_t + \varepsilon_{it} \qquad (2.4)
$$

should be biased upward relative to the IV regression.

2.4 Data

I build my sample from the annual and monthly surveys of the Townsend Thai project. In May, 1997 the project surveyed over two thousand rural households in four provinces. The annual survey followed the households from one-third of the baseline districts up through 2010 (Townsend et al., 1997). The monthly survey followed the baseline households plus several new additions from four of the remaining districts (Townsend, 2012). The monthly survey reocrds changes in household income, crop conditions, and many other features of the household. To these features I add the monthly international price of rice from January 1980 to June 2012, taken from the IMF's commodity price dataset.[8]

I use the monthly data to test the theory of risky income. My final sample contains all 743 households that responded to at least two of the seventy-two monthly rounds the project has released. Table 2.2 summarizes the sample characteristics. I observe the average household for 65 months, but have the full five years of data for over three-quarters of households. I mark a household to be a rice farmer if it harvests rice at any point in the sample. I mark a household as expecting a harvest if it harvests rice in the next three months; I mark it as having had a recent harvest if it harvested rice in the current month or the previous three months.[9] Table 2.2 shows that households expected a harvest one-fifth of the time. I define the number of economic activities as the sum of the number of "large" businesses, crop-plots cultivated, types of livestock raised, number of jobs held by all members, number of miscellaneous or small businesses, and an indicator for whether the household engages in aquaculture (raising fish or shrimp). I define total revenue as the sum of revenue from each economic activity. I define total consumption as all weekly and monthly household expenditure. Net transfers, which I use to classify households as insured in Section 2.3.1, are the total incoming transfers minus total outgoing transfers. I deflate revenue, consumption, and transfers to be in May 2007 Thai baht.[10]

Table 2.2 shows that the average monthly revenue is 620 U.S. dollars per month at May 2007 exchange rates. This figure is skewed upward because revenue is bounded below by zero but spikes during rice harvests; hence the high standard deviation. Consumption is less seasonal and the mean of 194 dollars is less skewed.

I test the theory of lumpy investments with the annual panel. In addition to the four provinces and roughly 1000 households followed from baseline, the project added two more provinces and roughly 500 more households several years into the survey (both from the new provinces and from the original villages to counter attrition). My final annual sample for the lumpy investment tests is 1502 households. I construct the number of activities as closely as possible to my monthly measure: the sum of the number of large businesses, crop-plots, jobs, herds, an indicator for aquaculture, and a subset of the miscellaneous income sources. [11] The annual average of 4.6 activities is almost identical to the monthly

[8] I treat a household-month surveyed in the first half of the month as though observed in the previous month when I merge with time series data and define month dummies. Since the rice price and consumer price index are monthly averages, my convention best matches the survey response period to the horizons of the aggregate prices.

[9] Some fraction of households claim to sell rice during months when they still expect a harvest. In Appendix B.4 I show that dropping these observations does not change the results.

[10] For more details on how I construct the variables, see Appendix B.2.

[11] Miscellaneous income sources in the annual survey often include remittances and other sources that do

Table 2.2
Descriptives of the Monthly Sample

Household-Month Mean and Standard Deviation				Fraction of Households or Household-Months	
Number of activities:	4.6 (3.3)	Revenue:	21352.8 (79854.7)	Rice Farmers:	0.48
				Of Whom	
Household size:	5.3 (2.4)	Consumption:	6692.2 (24449.2)	Fraction of time expecting harvest:	0.23
Total Labor:	80.0 (75.6)	Net Transfers In:	667.0 (35274.8)	Fraction of time just had harvest:	0.31
Households:	743	Avg. Obs/HH:	65.0	Observations:	48329

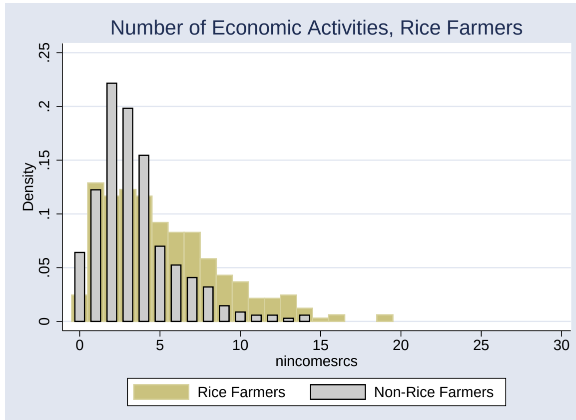
average in Table 2.2, but it varies less because the annual measure wipes out within-year variation in activities.

The histogram in Figure 2.6, which shows the distribution of number of activities in an arbitrary month, confirms that households in Thailand have many economic activities. Rice farmers are particularly under-specialized. Figure 2.7 graphs the top seven spontaneous responses to "What did your household do in the worst year [for income] of the last five to get by?" The most popular response was to take on an extra occupation, followed by working harder than usual. These responses do not prove households avoid risk through under-specialization, only that they cope with shocks through under-specialization. But if households must smooth their consumption by working harder, then they must have no better option. Borrowing money is only the third most popular response and using savings only the fifth. The fourth most popular response is to consume less, meaning many households lack even second-rate insurance.

Figure 2.8 shows the correlation between revenue and consumption, which direct evidence of imperfect insurance. I compute the correlation between monthly revenue and consumption expenditure for each household over however many months I observe it (72 months for the majority). If a risk-averse household has perfect insurance, its consumption should be independent of its current revenue; in fact, it should be constant. A household without perfect insurance cuts consumption when revenue falls, making the correlation positive. A higher correlation is evidence of less insurance. The figure plots the density

not meet my definition of economic activities (namely, revenue generating activities that require labor). I filter these unwanted sources using regular expressions on the textual descriptions of sources. The 1999 survey unfortunately does not contain textual descriptions, but the year dummies in the annual regressions should account for any 1999-specific measurement error.

Figure 2.6
Number of Economic Activities, Rice Farmers and Non-Rice Farmers



Note: The histogram depicts fraction of households with any number of economic activities in an arbitrarily chosen month. Rice farmers are more likely to have many activities.

of the correlation among rice farmers and non-rice farmers. Since zero is modal it appears many households do have near-perfect insurance, but many more do not. The distribution is heavily skewed towards less insurance with rice farmers particularly uninsured.[12] Some households have a negative correlation because of sampling error: the true correlation might be zero, but my estimate fluctuates around the truth and lands below zero for some households.

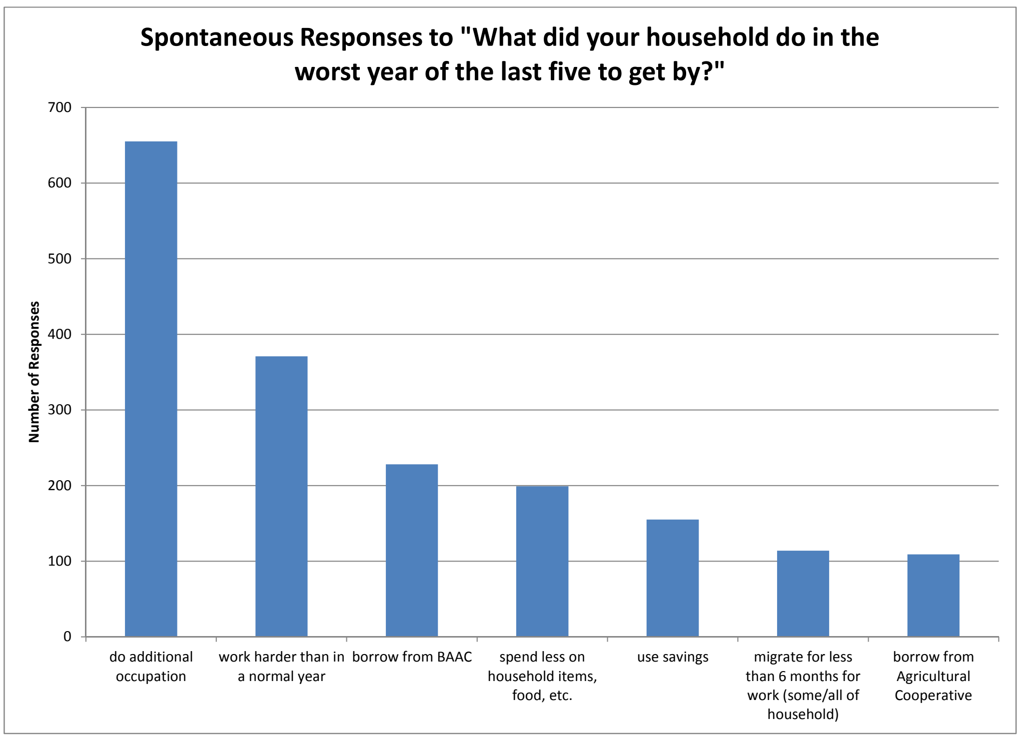
2.5 Risk and Under-Specialization Results

2.5.1 Main Results

Table 2.3 reports the results of the four tests derived in Section 2.2 and implemented as described in Section 2.3. Column 1 and 2 estimate (2.1) first with and without month dummies. Column 3 estimates (2.4), and Column 4 estimates (2.3). Aside from the ordinary least squares regression reported in Column 3, all regressions use the generated measure of volatility. I calculate the p-values and confidence intervals of these regressions using a two-stage bootstrap. The bootstrap, which I describe in detail in the online appendix,

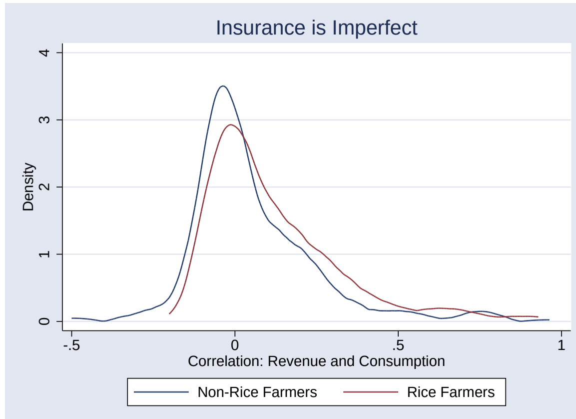
[12]The result may seem at odds with the high degree of insurance Townsend (1994) finds, but recall his result is that household consumption moves only with village-level and not household-level income. Figure 2.8 does not control for village-level shocks because a household cares only about having stable consumption, not where instability comes from. The shock I use for identification in Section 2.3 is a village-level shock: the international price of rice. It is precisely the village's inability to hedge against the price that drives households to under-specialize.

Figure 2.7
Household Response to Negative Income Shocks



Note: The 1997 round of the Townsend Thai annual survey asks households how they coped during the the worst income year of the last five. They first gave spontaneous responses, which the project classified into categories. The graph reports the frequencies of the seven most popular responses. Many households work more or spend less to absorb income shocks rather than borrowing or using savings.

Figure 2.8
Correlation Between Monthly Revenue and Consumption



Note: I compute each household's monthly correlation between total consumption and total revenue and plot the density of the correlation for rice farmers and non-rice farmers. Perfect insurance (whether self-insurance or otherwise) implies zero correlation. Almost all households have a positive correlation, meaning they absorb shocks to revenue by consuming less. The problem is worse for rice farmers.

corrects for the generated volatility measure and within-household correlation in the error term across time.[13]

The model's first test—Test 1—states that greater risk causes entry into more activities. The effect of risk on activities is the coefficient on $[ExpectingHarvest] \times [Volatility]$ in Column 1 of Table 2.3, and as predicted it is positive and significant. The model also predicts in Test 2 that higher expected returns to the primary activity (rice farming) should cause a decrease in activities. The coefficient on $[ExpectingHarvest] \times [Mean]$ confirms that higher returns have a negative and significant effect on number of activities. Column 2 verifies that both results hold when I include month fixed-effects, though the estimate of the effect of risk on activities becomes smaller.

Test 3 states that if the extra activities rice farmers take on cause their (side) revenue to fall, then the failure to specialize is costly. I implement the test by running using the regression in Column 2 as a first-stage regression for (2.3) using the response of farmers expecting a harvest to expected price and volatility as instruments for the number of activities. Column 4 of Table 2.3 reports the two-stage least squares coefficient on $[Activities]$ is negative and significant, confirming that under-specialization is costly. Column 3 reports the results of the simple ordinary least squares regression with month dummies of revenue

[13]It is not clear how to bootstrap the F-statistic on the excluded instruments or the Hansen's J Statistic. However, I can simply replace the generated volatility with the $|P_2 P_1|$ in the first stage. Since this is perfectly collinear with the generated measure it produces algebraically identical coefficients, but since it is not generated the standard F and J statistics are valid.

Table 2.3
Main Results

	(1) Activities	(2) Activities	(3) Revenue	(4) Revenue
Activities			1851.26***	-13883.30**
			[0.000]	[0.035]
Mean	-0.00*			
	[0.096]			
Volatility	-0.08***			
	[0.000]			
Rice Farmer				
- × Mean	0.01***	0.00		-128.97
	[0.010]	[0.359]		[0.135]
- × Volatility	-0.20***	-0.10***		-272.90
	[0.002]	[0.009]		[0.618]
Expecting Harvest				
- Main	1.82***	1.89***		4993.52
	[0.006]	[0.000]		[0.310]
- × Mean	-0.02***	-0.02***		(Exc. Inst.)
	[0.000]	[0.000]		
- × Volatility	0.18***	0.05*		(Exc. Inst.)
	[0.001]	[0.089]		
Recent Harvest				
- Main	-0.76	-0.57		-34753.04**
	[0.475]	[0.303]		[0.034]
- × Mean	-0.03***	-0.01***		300.00
	[0.000]	[0.002]		[0.129]
- × Volatility	0.41***	0.17***		234.48
	[0.002]	[0.004]		[0.819]
Household Fixed-Effects	Yes	Yes	No	Yes
Month Fixed-Effects	No	Yes	Yes	Yes
F-Stat Exc. Inst.				13.604
Hansen's J Stat.				0.125
Households	743	743	743	743
Observations	48329	48329	48329	48329

Note: The regressions the four tests of the theory of risky income (see Section 2.2): Test 1 (Risk): risk increases the number of activities; Test 2 (Returns): higher returns decrease the number of activities; Test 3 (Cost): more activities may cause side revneue to fall; Test 4 (OLS Bias): OLS is biased upwards. Column 1 estimates Equation 2.1, Column 2 estimates Equation 2.2, Column 3 estimates Equation 2.4, and Column 4 estimates Equation 2.3. The bracketed values are p-values. I compute the p-values in Columns 1,2 and 4 using a two-stage bootstrap that corrects for generated regressors and clusters at the household level (see Appendix B.3). I compute the p-values in Column 3 using the usual asymptotic standard errors that cluster at the household level. The value of the F-statistic on the excluded instruments from the first stage meets common standards for strength. The value of the J-statistic for overidentification is much too small to reject the null of exogenous instruments.

on number of activities (2.4). Test 4 states that the ordinary least squares coefficient on $[Activities]$ should be biased positively relative to the two-stage least squares coefficient because the farmers who pay lower costs for additional activities are exactly those who select into more activities. The coefficient is biased so strongly the sign flips, making under-specialization appear efficient.

What do the sizes of these coefficients mean? Since the average price volatility for all available months is 8.8, the regression in Column 1 implies a 10 percent rise in volatility causes the farmer to enter $.18/(1/8.8) * 10/100 = .16$ additional activities. A similar calculation shows that the more conservative estimate in Column 2 implies a 10 percent rise in volatility causes the farmer to enter .04 activities. Recall from Table 2.1, however, that the international price of rice is not perfectly correlated with the actual price the farmer receives. This may be because government price supports give the farmer some insurance. Regardless of the cause, since the international price has a correlation coefficient of roughly 1/3, a one unit rise in the volatility of the international rice price predicts a 1/3 unit rise in the volatility of the price the farmer receives. We can adjust the earlier numbers by dividing by 1/3, yielding estimates of .48 and .13 for the baseline and conservative estimates. The baseline estimate suggests the household enters an additional activity when the prices it faces become 21 percent more volatile.

The two-stage least squares estimate implies a household will forego over 13 thousand baht, or over 60 percent of its average monthly revenue, in any year. According to the model in Section 2.2 this estimate is actually biased upward, suggesting the true cost is even higher than implied. But recall that the average household has a little over four activities at once, making an additional activity a very large increase. Further, if the cost of an activity varies across households the estimate is not the average cost. If there is an upper bound on the number of activities a household can juggle, then the households with fewer activities are those most likely to respond to the instruments. These are also the households for whom an extra activity is most costly. Then the estimate, which is the continuous equivalent of the local average treatment effect, might be higher than the average cost of a side activity.

The responses of households who had a recent harvest bear some explanation. First, the coefficient on $[RecentHarvest] \times [Mean]$ is negative. Since the expected price after the harvest is correlated with the price received at harvest, the negative coefficient confirms both Figure 2.7 and the results of Adhvaryu et al. (2013), both of which say that households increase their number economic activities in response to bad income shocks. Finally, the positive and significant coefficient on $[RecentHarvest] \times [Volatility]$ seems puzzling, as risk should not matter after a household has had its harvest. There are two explanations for this. First, since the current volatility is correlated with past volatility, this may just reflect that the household faced risk before the harvest and took on extra activities. Since the household cannot drop the extra activities immediately after harvest—temporary jobs must be finished and small businesses must be wound down—the household may still have more activities than usual after harvest. The second possibility is that since current volatility simply says the price has moved drastically in the recent past. Since the expected price ([Mean]) does not perfectly capture the price at harvest, a high volatility means it is more likely the household had a low price at harvest. Since households take on activities to recover from low prices, the coefficient on post-harvest volatility may be picking up the response to negative income shocks.

2.5.2 Robustness

Table 2.4 reports several robustness checks. The theory in Section 2.2 assumes the total labor supplied by the household is fixed, but in truth the household may work less when the returns to its labor grow riskier. Alternatively, the household might send some members to work abroad or in Bangkok. Columns 1 and 2 show that the effects of higher volatility and higher returns on the number of activities remain unchanged when I control for the household's total labor and the number of household members. Likewise, Column 6 shows that the effect of additional activities on revenue remains unchanged.

Columns 3 and 7 of Table 2.4 both answer a simple concern: should we believe Thai rice farmers use a model of autoregressive conditional heteroskedasticity to decide how to spend their time? The model only formalizes a simple intuition: when prices fluctuate they are risky. Columns 3 and 7 confirm that simpler measures of the mean and volatility—the current price and the absolute value of the change in the price since last month—do not much change the results.

If the volatility of the price is just a proxy for unexpected decreases in the price, then what I assume is a response to risk may in truth be a response to changes in the household's permanent income. If this story is true, then the household should respond more strongly to simple changes in the price than to my measure of volatility, which is proportional to the absolute value of the change. Column 4 of Table 2.4 runs a regression that replaces my measure of volatility with the simple change in price. Households expecting a harvest do not respond to simple changes in the price.

The reader may also worry whether the expected price and the volatility are valid instruments for side revenue. If the price of rice and the price of corn, say, are correlated then the expected price is no longer a valid instrument for rice farmers who also grow corn. Column 5 of Table 2.4 verifies that the second stage results still hold when I use a measure

Table 2.4
Robustness: Household Labor and Size

	(1) Activities (Labor)	(2) Activities (HH Size)	(3) Activities (Non-Gen)	(4) Activities (Price Change)	(5) Revenue (Exc. Crop Rev.)	(6) Revenue (Labor & Size)	(7) Revenue (Non-Gen)
Activities					-11232.10**	-17060.61***	-9305.78*
					[0.031]	[0.009]	[0.064]
Rice Farmer							
- × Mean	0.00	0.00	0.01***	0.00	-112.59	-91.88	-21.25
	[0.249]	[0.441]	[0.001]	[0.854]	[0.178]	[0.278]	[0.795]
- × Volatility	-0.10***	-0.10**	-0.05***	0.02***	-493.25	-791.17	-397.12
	[0.006]	[0.013]	[0.000]	[0.000]	[0.378]	[0.265]	[0.120]
Expecting Harvest							
- Main	1.78***	1.86***	2.58***	2.09***	7295.65	4897.58	1537.60
	[0.000]	[0.000]	[0.000]	[0.000]	[0.118]	[0.345]	[0.725]
- × Mean	-0.01***	-0.01***	-0.02***	-0.01***	(Exc. Inst.)	(Exc. Inst.)	(Exc. Inst.)
	[0.000]	[0.000]	[0.000]	[0.000]			
- × Volatility	0.05*	0.05*	0.02***	-0.00	(Exc. Inst.)	(Exc. Inst.)	(Exc. Inst.)
	[0.094]	[0.088]	[0.003]	[0.881]			
Recent Harvest							
- Main	-0.54	-0.61	1.30***	0.24	-28550.62*	-34442.87**	148.02
	[0.306]	[0.281]	[0.000]	[0.423]	[0.052]	[0.030]	[0.992]
- × Mean	-0.01***	-0.01***	-0.02***	-0.01*	167.55	288.46	-44.62
	[0.002]	[0.005]	[0.000]	[0.062]	[0.381]	[0.129]	[0.758]
- × Volatility	0.16***	0.17***	0.05***	-0.04***	1018.61	161.63	965.18**
	[0.003]	[0.002]	[0.000]	[0.000]	[0.390]	[0.878]	[0.012]
Household Labor	0.01***					275.42***	
	[0.000]					[0.000]	
Household Size		0.13***				950.03	
		[0.000]				[0.312]	
Household Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat Exc. Inst.					13.604	13.154	20.660
Hansen's J Stat.					0.018	0.273	0.369
Households	743	743	743	743	743	743	743
Observations	48160	48164	48329	48329	48329	48160	48329

Note: The household might work more or recall migrants from cities in response to higher volatility. The extra workers might take up new activities without actually affecting the degree of specialization. In Columns 1 and 2 I control for each and show the results of the main text are unchanged. (Controlling for both simultaneously has similar effects.) In Column 3 I verify that controlling for household labor and household size does not qualitatively change the second-stage results.

Figure 2.9

Households Receive More Transfers when Prices are Low at Harvest

Note: The first bar depicts average incoming transfers for households harvesting rice when the international rice price is "normal"—above the bottom quartile of all prices I observe in the period covered by the monthly panel. The second bar depicts the average transfers when prices are "low"—in the bottom quartile. Rice farmers receive more money when the value of their harvest is low.

of revenue that excludes earnings from crops. In Appendix B.4 I show that median wages in each village are not affected by average movements in the regressor of interest.

The final question is whether what I measure is really a response to risk. To test this I examine whether households with better insurance respond less to changes in the volatility. In poor countries a household often relies on family and friends for support in hard times. [14] Figure 2.9 shows the rice farmers in my sample are no different: when the international price is low rice farmers tend to receive more transfers. I calculate for each household the monthly correlation between its net incoming transfers and its revenue, and call a household "insured" if that correlation is negative.

Columns 1 and 2 of Table 2.5 report the separate responses of the uninsured and insured sample. As expected, the response of the insured is smaller and insignificant. Since my measure of insurance is not exogenous I cannot rule out that households with insurance differ from uninsured households in ways that change how they might respond to volatility. Still, though not a perfect test of the model the result is consistent with the model.

[14] Rosenzweig (1988) found that households structure themselves to ease income sharing. Townsend (1994) and more recently Munshi and Rosenzweig (2009) find village and caste networks provide insurance in India. Yang and Choi (2007) show that rural Filipino households who suffer bad rainfall shocks receive more remittances from overseas family.

Table 2.5
Check: Insurance

	(1) Activities (Uninsured)	(2) Activities (Insured)
Rice Farmer		
- × Mean	0.01**	-0.00
	[0.027]	[0.922]
- × Volatility	-0.09*	-0.10**
	[0.069]	[0.013]
Expecting Harvest		
- Main	1.65***	1.99***
	[0.001]	[0.000]
- × Mean	-0.02***	-0.02***
	[0.000]	[0.000]
- × Volatility	0.09*	0.04
	[0.067]	[0.277]
Recent Harvest		
- Main	-0.77	-0.39
	[0.186]	[0.557]
- × Mean	-0.01	-0.01***
	[0.221]	[0.006]
- × Volatility	0.14**	0.19***
	[0.025]	[0.003]
Household Fixed-Effects	Yes	Yes
Month Fixed-Effects	Yes	Yes
Households	270	473
Observations	16933	31396

Note: I split the sample into households who receive transfers of income when their consumption is low ("insured") and those that do not ("uninsured"). I confirm that volatility has a larger effect on the number of activities among households that are uninsured.

2.6 The Alternative Theory: Lumpy Investments

If "the poor cannot raise the capital they would need to run a business that would occupy them fully" (Banerjee and Duflo, 2007) then poor households cannot specialize. Suppose a man can learn to sew or bake but cannot can sew more than a few shirts or bake more than a few loaves without a sewing machine or an oven. Since he cannot afford either investment he cannot grow either business and cannot support his family unless he sells both shirts and bread. This is the theory of lumpy investments.[15]

To test the theory I exploit a government program that produced quasi-experimental variation in the supply of credit. The theory predicts that households that get more credit should be better able to make the lumpy investments that let them specialize. The Million Baht Program gave one million baht to a fund for public lending in every village in my sample. Kaboski and Townsend (2009, 2011), who are the first to exploit the program, argue that the boundaries of villages in Thailand are set by bureaucratic fiat rather than economic logic. The sizes of villages are effectively random. Since every village got the same amount of credit the per-household increase in credit is also random, with smaller villages exogenously given more credit. Kaboski and Townsend (2011) confirm in their first table that small and large villages have parallel trends.

Since I do not know when in 2001 the program reached each village, I use the annual data and treat 2001 as the year of implementation. The effect of the program is measured by the interaction of the year of implementation interacted with some measure of village size. In one specification I use an indicator for whether the household is in the bottom quartile in number of households; in the other I use the actual per-household injection (1 million/number of households).[16] The theory predicts that the signs of the coefficients should be negative and significant.

According to Table 2.6 the coefficients are insignificant and have the wrong sign. The positive coefficients on $2001 \times Small$ and the other interactions are not consistent with the lumpy investment theory, but might be consistent with the model from Section 2.2. If risk is really what drives under-specialization and some households want more activities but cannot afford to pay the fixed-cost, giving them credit might let them enter more ac-

[15]The inability to invest may create another source of under-specialization: the need to take on extra jobs because one may only work so long at any single task. Suppose labor and capital are complements, and make it simple with an extreme example: perfect complementarity. Suppose an activity m produces revenue with production function $y^m = A^m \min[L, K]$, with $m = T, B$ for tailoring or baking. Suppose $A^T > A^B$ for some household. If the household's labor endowment is $\bar{L}$, it will specialize in tailoring with $K^* = \bar{L}$. But suppose increasing capital beyond $\tilde{K} < K^*$ requires a lumpy investment the household cannot afford. If the household specialized, it would be left with $\bar{L} - \tilde{K}$ units of unused labor. In other words, it would be idle. The alternative is to spend its remaining time baking, so its total revenue is $A^T \tilde{K} + A^B(\bar{L} - \tilde{K}) < A^T \bar{L}$.

[16]The results do not change when I use the log of the injection as in Shenoy (2014a).

Table 2.6
Testing the Theory of Lumpy Investments

	(1) Activities b/se	(2) Activities b/se	(3) Activities b/se	(4) Activities b/se
Small Village	-0.010		0.102	
	(0.11)		(0.11)	
2001 X Small	0.132		0.175	
	(0.16)		(0.16)	
2002 X Small	0.213		0.144	
	(0.14)		(0.15)	
Credit/HH		3.010		9.977
		(13.47)		(13.28)
2001 X Credit/HH		11.540		11.569
		(8.63)		(8.20)
2002 X Credit/HH		21.857*		16.619
		(11.55)		(11.86)
Household Fixed-Effects	Yes	Yes	Yes	Yes
Month Fixed-Effects	Yes	Yes	Yes	Yes
Villages	80	80	64	64
Households	1502	1502	706	706
Observations	15340	15340	9884	9884

All standard errors clustered by village

Note: The regressions test the lumpy investment theory using the Million Baht Program. The coefficient on the interaction of village size with the year of implementation (2001) estimates the effect of relaxed credit on number of activities. The measure of number of activities is similar as possible to that in the risk regressions. A village in the bottom quartile of number of households is “small”. The alternative specification uses the average per-household credit injection (one million divided by number of households). The first two columns use the largest possible sample of households while the last two use a balanced panel. The lumpy investment theory predicts the program’s impact should be negative and significant, which it is not. All inference uses the usual asymptotic standard errors that cluster at the village-level.

tivities. But this story, which lacks direct evidence and rests on coefficients that are not signficant, remains only a story. Only the coefficient on $2002 \times Credit/HH$ in Column 2 is (marginally) significant, and even that significance vanishes when I restrict the estimation in Column 4 to a balanced panel.

My results do not mean credit constraints have no effect on specialization. Aside from the usual caveats—a lack of evidence is not a rejection, and rejection in Thailand does not mean rejection in other countries—I only test a limited form of the theory. The smallest villages received per-household credit injections of half the median income. If households need sewing machines and ovens the credit injection would cover it. Since most micro credit charities believe small entrepreurs need small loans, finding no effect from a small rise in the supply of credit is not trivial. But if a few households want to build factories that would provide stable and salaried jobs to everyone else, the Million Baht Program is too small.

2.7 Summary

I show that Thai rice farmers expecting a harvest increase their number of economic activities when confronted with more volatile prices. My estimates suggest a 21 percent rise in volatility causes a household to enter an extra activity. I use this exogenous change in the number of activities to verify under-specialization reduces revenue. Finally, I test an alternative theory of under-specialization—that the poor run many small businesses because they cannot afford the lumpy investments needed to grow any one—and find no supporting evidence.

The pin-maker wastes time when he switches from straightening wires to cutting them, and I find evidence of this waste in rural Thailand. My results do not measure the talent and investment wasted when the poor forego expertise in a single trade or investment in a single business. This kind of under-specialization, which changes the structure of an economy, is a long-run cost that requires a long-run study. Future research must test whether long-run risk causes long-run under-specialization and how much it costs.

CHAPTER 3

Does Factionalism Distort Production? The Case of Caste Politics in Rural India

3.1 Introduction

It is sometimes deemed acceptable for governments to discriminate. They charge higher taxes to the rich, give cheap loans to students, and give subsidies to farmers without (much) outcry. But when governments discriminate for political rather than economic reasons they are accused of factionalism. Favoring the supporters of those in power is a mark of misgovernance. Using the term “extractive institutions,” Acemoglu and Robinson (2012) blame such misgovernance for many countries’ underdevelopment. By distorting production factionalism keeps poor countries poor. Or at least that is the prevailing belief.

This paper shows that the prevailing belief is not always true. I build a model in which firms are aligned either with or against the faction in government. Firms can increase their productivity if they get a government service. The bureaucrat who provides the service can demand a bribe, and he is less likely to be caught if he demands bribes from firms in the losing faction. He responds by charging fewer bribes to winners; this is factionalism. But the bureaucrat sets his bribe like a monopolist sets a nonlinear price; he maximizes economic surplus and then extracts it all. Since the bribe must not distort the firm’s behavior it must not depend on the firm’s choice of inputs. It depends only on faction. Ironically the very fact that the bureaucrat discriminates for political rather than economic reasons prevents any distortion.

I test the model by studying how the outcome of a village council election affects Indian farmers. In India the services of both the national and local government flow through local bureaucrats. A bureaucrat who grants ration cards or land titles has a chance to extract bribes. Since the village council and its president can fire the bureaucrat he risks his job if he demands bribes of the president’s faction. Factions in an Indian village are typically based on caste. Since each farmer’s caste is known the bureaucrat can tell which are aligned with the president. A clear test for factionalism is to check whether being of the same caste as the council president causes a household to pay fewer bribes.

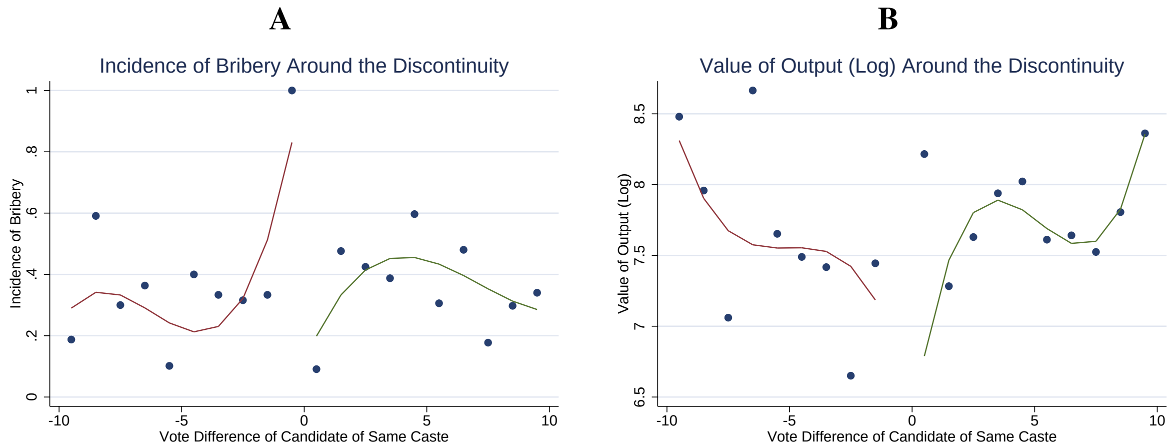
I implement the test using a regression discontinuity design. I define the vote difference of each candidate as the difference between her share of the vote and the share of whichever other candidate got the most votes. For each farmer I define the running variable as the vote difference of the candidate with the same caste. When this vote difference switches from negative to positive the household's alignment switches from the loser to the winner. Households just above and below the threshold are otherwise identical. By checking whether the probability the household pays a bribe changes at the discontinuity I test for factionalism. By next checking whether farm output changes I test whether factionalism distorts production.

I combine data on bribes, farm production, and election outcomes from a panel of rural Indian households. For each administration of the village council, households report which officials they approached for government services and of these which demanded a bribe. Households also report their farm production in one year of the same administration. Finally, households report their caste and subcaste. Using these reports I link households to candidates who contested the election that determined each administration.

Figures 3.1 summarizes my results. I look at two outcomes: a dummy for whether the household paid any bribes, and the log of the value of farm output. I regress each on a cubic polynomial in the running variable (with coefficients that may differ on either side of the discontinuity), and an indicator for whether the household's candidate won. Panel A of Figure 3.1 shows that in approaching the discontinuity the probability of having to bribe an official rises among losers and falls among winners. After a close election bureaucrats are far less likely to demand bribes from households of the same caste as the winner. But according to Panel B of Figure 3.1 this evidence of factionalism does not come with any evidence of a distortion to production. The log of the value of farm output falls near the discontinuity for both winners and losers. I run the same test on measures of input use—land sown, fertilizer used, pesticides bought—and find no evidence that farmers in the losing faction use less. As the model predicts, I find no evidence to suggest factionalism distorts whether a farmer seeks government services. Losing households borrow no less money and pay no more interest.

This paper adds to three areas of literature. The first explores the causes and costs of corruption. Banerjee (1997) creates a model where bribery and red tape arise from a conflict of interest between a benevolent government and the self-interested bureaucrats who implement its policies. Acemoglu and Verdier (2000) build a model with a similar conflict of interest to show that the ultimate cause of a failure of governance may be a

Figure 3.1
Panel A: Probability of Paying a Bribe; **Panel B:** Log of the Value of Farm Output



Note: The panels plot the predicted values of the incidence of bribery and the log of the value of farm output based on a cubic polynomial in the vote difference. The coefficients of the polynomial can vary on either side of zero. The dots are the average bribe paid within a ± 0.5 percentage point bin. Each bin may contain a different number of households and elections.

failure of markets. Like Banerjee's model mine predicts an efficient allocation despite corruption, but the reason is different. His model assumes the benevolent government can detect misallocation and punish bureaucrats for it. My model assumes the government is actually hostile to some firms. Yet bureaucrats still want the efficient outcome because it maximizes their profits. This alignment of incentives also explains why the prediction of my model differs from that of Acemoglu and Verdier. In both models the bureaucrat sells each firm the outcome it wants, but the market failure in their model makes each firm want an inefficient outcome.

Another literature studies how ethnic conflict distorts a government's ability to provide ethnic goods. Miguel and Gugerty (2005) show that ethnically diverse villages in Kenya provide fewer public goods, leaving everyone worse off. Figure 3.1.B is consistent with their result; the output of winners and losers alike falls after the more polarized electons near the discontinuity. But Miguel and Gugerty do not study whether winners and losers are treated differently. Besley et al. (2004) comes closest to my question and context. They show that low caste Indian households are more likely to receive government benefits in villages reserved for low caste candidates. But villages may not be randomly chosen for reservation in the states they study. It is not clear whether the effect they find is causal. My study can identify the causal effect of having a president of the same caste. Moreover, I study the effect on production rather than the distribution of government benefits.[1]

[1] The government services that matter most to farmers are non-rival whereas ration cards and other such

My paper links these two literatures to a third that studies the misallocation of productive resources. Banerjee and Munshi (2004) show that ethnic divisions cause misallocation in the garment district of Thirapur, but do not study political power. Fisman (2001) shows that having ties to Suharto's regime affects the prices of stocks in Indonesian firms, but cannot measure the direct effect on production. Hsieh and Klenow (2009) show that state owned enterprises in China are allocated too much capital, but cannot claim the pattern is causal. Finally, I show in earlier work (Shenoy, 2014a) that factor and financial market failures among Thai rice farmers cause surprisingly little misallocation. Unlike my earlier work, this paper isolates one cause of market failures: factionalism. The results suggest the link from factionalism to misallocation is subtle. Even severe factionalism need not cause any misallocation.

3.2 Theory

Suppose each firm chooses a vector of inputs $\boldsymbol{X}$ and produces revenue using a production function $A_i f(\boldsymbol{X}; G)$. A_i is the firm's productivity and $G \in \{0, 1\}$ is an indicator for whether the firm has access to a government service. For clarity I assume the service is binary, but the result is unchanged if G is continuious.

Assume f is strictly concave and

$$\frac{\partial f(\boldsymbol{X}; G = 1)}{\partial \boldsymbol{X}} > \frac{\partial f(\boldsymbol{X}; G = 0)}{\partial \boldsymbol{X}}$$

meaning the service increases the return to every input. The service might directly raise productivity, as irrigation or electrification would, or it might grant access to the credit the firm needs to make an investment.

The firm solves

$$\max_{\boldsymbol{X}} A_i f(\boldsymbol{X}; G) - \boldsymbol{w}^T \boldsymbol{X}$$

For each firm i the optimal choice of inputs $\boldsymbol{X}_i^*(\cdot)$ and optimal level of output $y_i^*(\cdot)$ depends on whether the firm has the government service:

$$\boldsymbol{X}_i^*(A_i; G = 1) > \boldsymbol{X}_i^*(A_i; G = 0)$$
$$y_i^*(A_i; G = 1) > y_i^*(A_i; G = 0)$$

benefits can only go to so many households. Even assuming Besley et al. do find a causal effect this difference in the services they study would create a difference in the effect of factionalism.

Then the benefit of the service to firm i is

$$b_i = y_i^*(G = 1) - y_i^*(G = 0)$$

The bureaucrat who provides the service earns a wage w, which he may supplement by demanding a bribe. The firm could respond to his demand by complaining to higher authorities. Only some firms have this option; for example, those with managers who are friends with a powerful politician. The firm is more likely to have friends in high places if it is a member of the faction in power—that is, if the firm's faction $F = W$ for winner rather than $F = L$ for loser. Let p_F be the probability a firm of faction F can complain. Then

$$p_W > p_L.$$

When a firm complains the bureaucrat is fired. He must return the bribe and loses his wage w. Given this risk the bureaucrat must decide whether to demand a bribe and if so how large the bribe should be.

Suppose the bureaucrat has decided to demand a bribe. He sets the size of the bribe using a function $B(A_i, \boldsymbol{X}_i, F_i)$. The $B(\cdot)$ that maximizes the bureaucrat's bribe is the same as the policy that maximizes a monopolist's profit: a nonlinear price. The bureaucrat maximizes economic surplus then sets a bribe that extracts it all:

$$B^*(A_i, \boldsymbol{X}_i, F_i) = b_i$$

The firm is left indifferent between taking and not taking the government service. By letting the bureaucrat set this bribe I have assumed he knows firm i's productivity A_i.

Now consider the bureaucrat's decision of whether to demand a bribe. He will demand a bribe if the expected benefit of getting away with it exceeds the certain benefit of asking no bribe and drawing the wage:

$$(1 - p_F)(b_i + w) > w \tag{3.1}$$

Define

$$\bar{b}_i = \frac{b_i}{A_i} = f[\boldsymbol{X}^*(G_F = 1); G_F = 1] - f[\boldsymbol{X}^*(G_F = 0); G_F = 0]$$

Suppose A_i has a distribution function H. Since factions are purely political H is the same for both the winning and losing faction. That is, $H(A_i \mid F = W) = H(A_i \mid F = L) = H(A_i)$. This assumption holds in my empirical work because I use a regression discontinuity to ensure the winning and losing faction are otherwise identical.

Rearrange (3.1) and apply the distribution function. The probability the bureaucrat asks for a bribe is

$$P_F = 1 - H\left[\left(\frac{1}{1-p_F} - 1\right)\frac{w}{\bar{b}_i}\right]$$

Clearly $P_W < P_L$; the bureaucrat is less likely to demand bribes of a firm in the winning faction. But the size of the bribe is chosen not to distort the firm's actions. It follows that all firms, whether winners or losers, will get the same number of services, choose the same level of inputs, and produce the same level of output:

$$\mathbb{E}[G_i \mid F = W] = \mathbb{E}[G_i \mid F = L]$$
$$\mathbb{E}[\boldsymbol{X}_i^* \mid F = W] = \mathbb{E}[\boldsymbol{X}_i^* \mid F = L]$$
$$\mathbb{E}[y_i^* \mid F = W] = \mathbb{E}[y_i^* \mid F = L]$$

3.3 Bribes and Factions in Rural India

The bureaucracy stands between India's public and its public goods. Land records, ration cards, guaranteed employment, and even a connection to the power grid come only with the signature of a bureaucrat, and his signature often comes only with a bribe. Corruption in India has inspired protests, legislation, and even Bollywood movies. All have failed to end it.

Using data from my sample, Figures 3.2 and 3.3 summarize which officials and services command the most bribes. Households were asked to reflect on the most recent administration of the village government. Each household reported which bureaucrats it approached for government services, which of these bureaucrats asked for bribes, and how big was each bribe.

Figure 3.2 shows the ten most corrupt bureaucrats based on what fraction of all bribes they accounted for. Together these ten account for over 70 percent of the incidence and over half of the value of all bribes. The Lekhpal, or land secretary, illustrates the power of the bureaucracy over the public. Since he issues the Record of Rights—the title to a piece of land—no one can prove they own their land without his help. Without such proof a bank will not accept land as collateral. Given the land secretary's importance it is not surprising most households who asked his help had to pay a bribe. The other prodigious bribe-takers are similar; they stand between households and an important government service. Figure

3.3 reports the ten services that account for most bribes. Getting the Record of Rights tops the list, but all these services unlock a government benefit or verify a household's condition.[2]

These bureaucrats answer to the Gram Panchayat, a council of locally elected representatives. The council is led by a president, the Pradhan, who in most states is chosen every five years by direct election. Though their exact powers vary by state, in all cases the president and council manage the local bureaucracy. In Karnataka the "Gram Panchayat may reduce in rank, remove or dismiss any employee," and the president has veto power over the hiring and wages of each bureaucrat (Govt. of Karnataka, 1993).

In some cases the president may explicitly tell bureaucrats to help friends and hurt foes, as in the village of Meerapur:

> Thirdly, Gaj Pal [the new council president] developed his nexus with the functionaries of the village including the Lekhpal. . . and sought their favours to extend support to some and withdraw the same from many (Sharma, 1988, p. 63).

But even without instruction the bureaucrat knows it unwise to demand bribes of the president's supporters. Though the bureaucrat may not know each villager's faction, in a small village he does know each villager's caste. As I show in Section 3.4, caste and faction are often the same.

3.4 Empirical Approach

Simply comparing how many bribes are paid by farmers aligned with the president to farmers aligned with her opponent is a flawed test of factionalism. The winners and losers of an election differ for reasons other than faction. Regardless of faction, bureaucrats will treat them differently. For example, suppose every election pits a rich faction against a poor faction. If there are more poor people than rich people the members of the losing faction will on average be richer. A bureaucrat who charges rich people more bribes will seem to favor winners over losers even if he cares nothing for faction. What looks like factionalism will in truth be only selection bias.

I handle selection bias by exploiting the regression discontinuity created by a close election. Let C be the set of candidates competing in a first-past-the-post election. Suppose Candidate A's share of the vote is s_A. Define her vote difference as

[2]Tables C.1 and C.2 in Appendix C.1 show the numbers used to construct these graphs. I compute these numbers from the reports of households in my sample (see Section 3.5). The data cover the period since the most recent village council election.

Figure 3.2
Officials Most Often Bribed

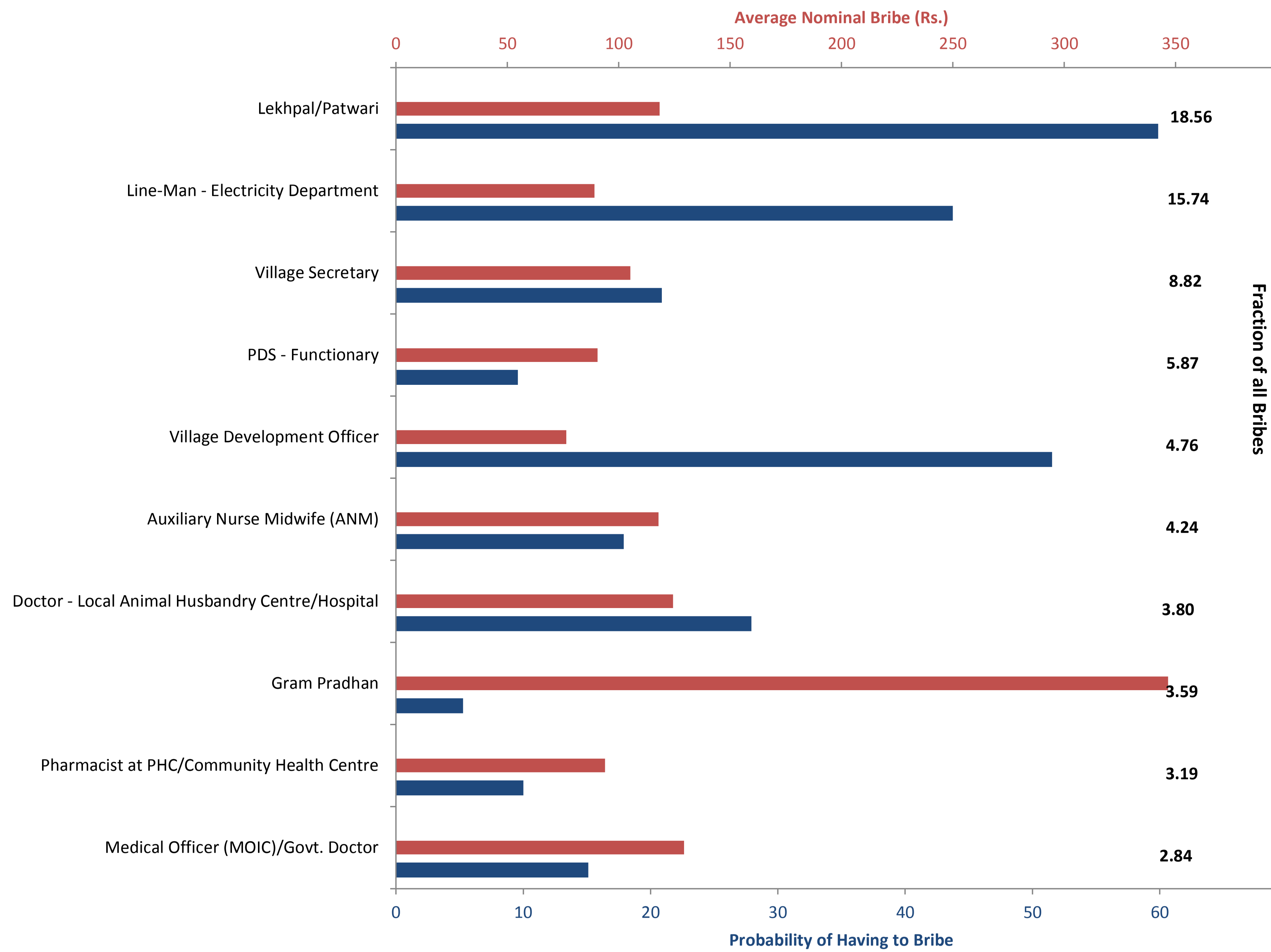


Figure 3.3
Services that Account for Most Bribes

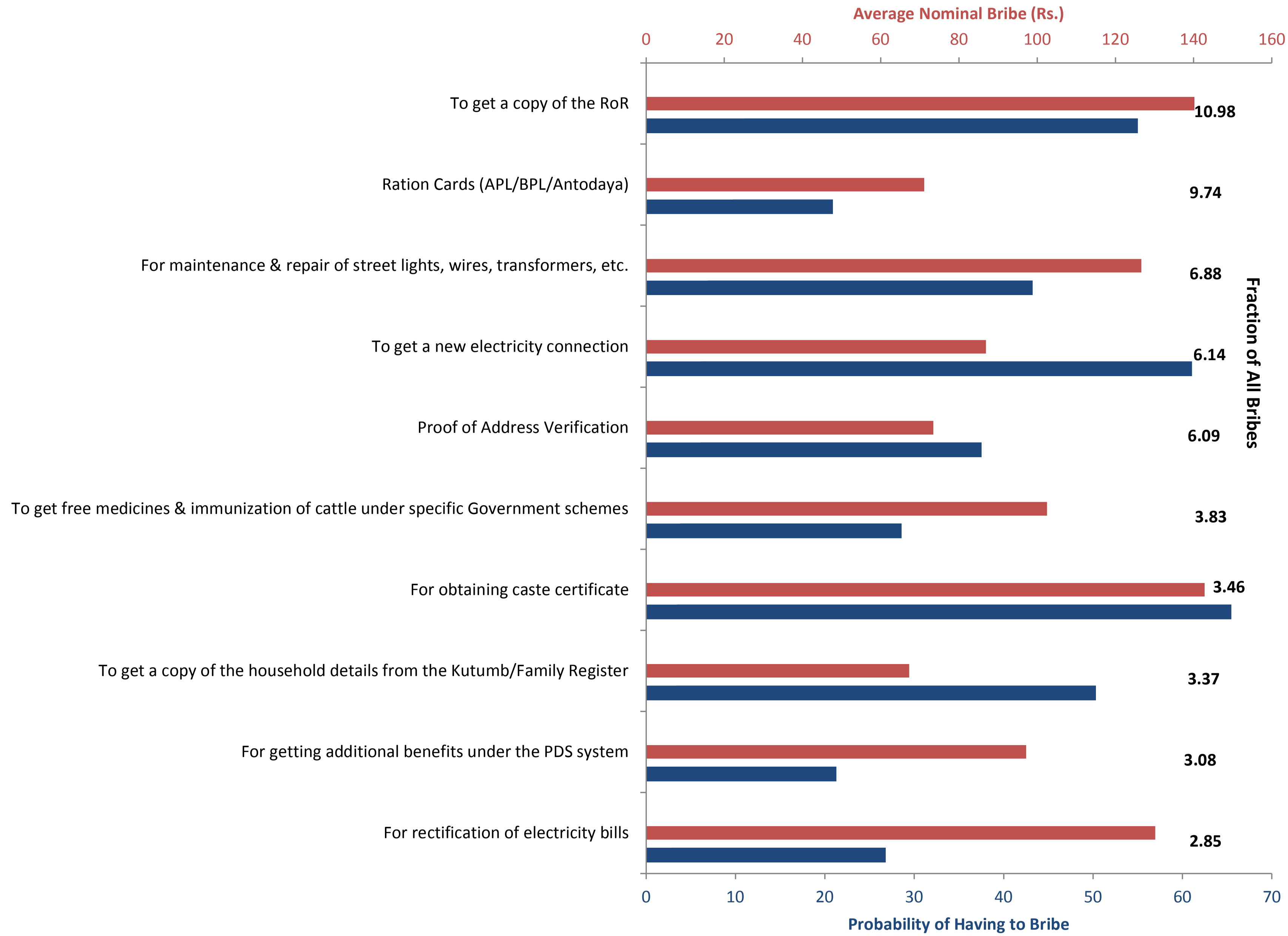
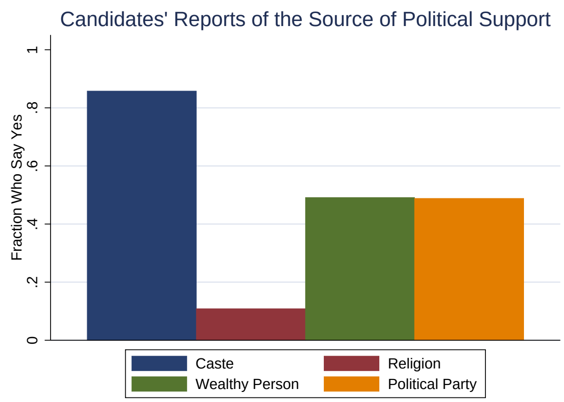


Figure 3.4
Candidates Rely on Caste Groups for Support



$$[Vote\ Difference] = s_A - \max_{c \in C \setminus \{A\}} \{s_c\},$$

the difference between her share of the vote and the share of the highest vote-getter excluding her. For the winner of the election the difference is the gap between her share and the share of the runner-up; for every other candidate it is the (negative) gap between their share and the share of the winner. The difference switches from negative to positive when a candidate switches from losing to winning the election.

Before I can find the effect of being in the losing faction I must first assign each faction a candidate and each farmer a faction. The survey from which I draw my household sample also collected data on each candidate who ran for the council presidency. For each candidate the survey asks whether the candidate relied on any of several sources of political support. Figure 3.4 compares what fraction of candidates relied on caste, religion, a wealthy person, or a political party. Caste is by far the most important; it is a natural basis for a faction. I assign to each farmer the vote difference of the candidate of the same sub-caste (uppajati). When the farmer's vote difference switches from negative to positive he switches from being in the losing faction to being in the winning faction. Farmers just above and just below the discontinuity should be similar in every dimension except that those above are in the winning faction. By comparing the bribes paid on either side of the discontinuity I can identify the causal effect of being in power on bribes paid.

My research design fails if farmers can manipulate their vote share. An individual household can only make such manipulations by changing its caste, which is impossible. But a caste group might manipulate a close election by consistently picking a better candidate or consistently rigging the outcome. Households of such a caste would differ from other households for reasons other than the election. I test for such differences by checking whether pre-determined outcomes—religion and caste, or wealth held before any election in my sample—change at the discontinuity. As I show in Section 3.6, there are no such changes.

Since nothing changes at the discontinuity except who wins the election, I can control for all other factors with a polynomial in the vote difference. As bribery may be a complicated function of the vote difference, I follow the literature by restricting my sample to a window around zero. The restriction reduces the number of twists and turns the polynomial must fit. I estimate a cubic polynomial within the window of vote differences between -10 and 10, and verify the main results do not change when I tweak the window and the degree of the polynomial. In all specifications I let the coefficients of the polynomial differ on either side of zero. I estimate a polynomial in a wider window around the discontinuity instead of estimating lines in a very narrow window as Imbens and Lemieux (2008) propose. The linear method is simpler but I lose elections (and thus clusters) too quickly as the window shrinks.

Let $\tilde{x}$ be the vote difference. I estimate

$$\begin{aligned}[Bribery] = \beta_0[Constant] + \alpha[Past\ Post] + \beta_1\tilde{x} + \beta_2\tilde{x}^2 + \beta_3\tilde{x}^3 \\ + [Past\ Post] \times (\beta_1'\tilde{x} + \beta_2'\tilde{x}^2 + \beta_3'\tilde{x}^3) + \varepsilon\end{aligned} \tag{3.2}$$

where $[Bribery]$ is some report of bribery and the indicator $[Past\ Post]$ switches on if the farmer's candidate won. Since the vote difference varies only between elections I cluster the standard errors by election. The coefficient on $[Past\ Post]$ measures how many fewer bribes a household must pay when its candidate wins.

After I establish the presence of factionalism I estimate

$$\begin{aligned}[Outcome] = \gamma_0[Constant] + \tau[Past\ Post] + \gamma_1\tilde{x} + \gamma_2\tilde{x}^2 + \gamma_3\tilde{x}^3 \\ + [Past\ Post] \times (\gamma_1'\tilde{x} + \gamma_2'\tilde{x}^2 + \gamma_3'\tilde{x}^3) + \varepsilon\end{aligned} \tag{3.3}$$

where $[Outcome]$ is access to a government service, the use of a farm input, or the output of the farm. If factionalism distorts production the coefficient on $[Past\ Post]$ will be significant.

3.5 Data

I construct my sample from the 2006 round of the Rural Economic Development Survey (REDS). These data were collected by India's National Council of Applied Economic Research in collaboration with several U.S. researchers. The first round of the survey was in 1969; later rounds in 1970, 1971, 1982, 1999, and 2006 resurveyed as many of the original households as possible.

The household module asked Indian villagers their caste and sub-caste, which are converted into a caste code. The village module records each candidate for the position of council president (Pradhan) and the number of votes she received. I discard tied elections, then for the remaining elections calculate for each candidate the share of the vote she received and the vote difference.

The survey does not record the castes of candidates but does record their names. In most Indian states a person's last name is simply his caste in the local language. For each village I build a crosswalk between caste and name by assigning to each last name the caste most frequently associated with it. [3] Since Indian names are often transliterated many different ways I use the user-written Stata command reclink to make a fuzzy match on last names between my crosswalk and the candidates. I discard any match made with less than 90 percent certainty. For each election I link households to the vote difference of the candidate with the same caste code; if a caste has more than one candidate I use the candidate who received the most votes. I drop households who cannot be linked to any candidate.

The survey also asked each household which village officials it approached for services, and of those which it had to bribe. I define an indicator for whether the household had to bribe of these, dropping households who did not approach any officials. I also calculate what fraction of officials each household bribed; if the household approached only the land secretary and the development officer, and had to bribe the land secretary, I code it as having bribed half of the officials approached. The survey asked households about the current council period and the previous two. By linking the bribes paid in a council period to the election preceding it I can examine the bribes paid as a function of the vote difference of the household's caste. The survey also asks which services the household sought from the government (see Figure 3.3 for examples), which I use to construct total services sought. I restrict my sample to households who cultivate land because my next step is to test whether factionalism distorts farm production.

The 2006 survey records the value of crops produced, (gross) land cultivated, and

[3] I throw away households who give only one name and names linked equally often to two different castes.

Table 3.1

Sample Descriptives

Fraction Bribed	0.11	Scheduled Caste/Tribe	0.07
	(0.20)		(0.25)
Gave Bribe	0.36	Muslim	0.03
	(0.48)		(0.16)
Vote Difference	2.88	Average Interest on Loans	6.04
	(22.90)		(11.92)
Past Post	0.65	Got Loan from Government	0.09
	(0.48)		(0.28)
Income, 1982 (Rs.)	3058.09	Value of Crops (Rs.)	3152.25
	(2463.01)		(4840.73)
Land, 1982 (Hect.)	3.99	Value of Chemical Fetilizer (Rs.)	239.90
	(3.85)		(398.67)
Full Sample		$-10 < Vote\ Difference < 10$	
-Observations	10017	-Observations	1054
-Elections	264	-Elections	88
		$-5 < Vote\ Difference < 5$	
		-Observations	420
		-Elections	37

Note: I calculate the means and standard deviations for the largest possible sample. Past Post is an indicator for whether the vote difference is positive, meaning the candidate of the household's caste won the election. Fraction Bribed is the fraction of officials approached for services that the household had to bribe. The value of output, the value of expenditure, and income are all deflated to 1960 rupees.

spending on chemical fertilizer, pesticides, and irrigation. But these numbers measure production only in the most recent election. To expand the sample I merge data on production, recorded in an earlier round of the same survey, from 1999. Using the dates of each election in each village I match the earlier data to the administration after either the previous or the previous-to-previous election. After merging I have data on production for 566 observations from 54 elections.

The 2006 survey also records each loan the household requested over the last few decades. For each household and each council period I compute the amount borrowed, the average interest rate weighted by loan size (conditional on borrowing anything), and whether the household received a loan from the government. I mark a loan as from the government if it came from a cooperative society, a nationalized bank, a rural development bank, a government scheme, or the Kisan credit program.

To test whether pre-determined variables change at the discontinuity I merge data on income, wealth, savings, land, the size of the household, and the age of the household head from 1982. I have data from 1982 for 738 observations from 82 elections. To the 1982 variables I add two fixed household characteristics, religion and caste, from the 2006 survey.

Table 3.2
Evidence of Factionalism

	A. Main Sample			
	Gave Bribe	Fraction Bribed	Bribe for Ration Card	Bribe for Rec. of Rights
Estimate	-0.937***	-0.381***	-0.312*	-0.442***
	(0.336)	(0.146)	(0.167)	(0.120)
Observations	1054	1054	1054	1054
Elections	88	88	88	88
	B. 1982 Sample			
	Gave Bribe	Fraction Bribed	Bribe for Ration Card	Bribe for Rec. of Rights
Estimate	-0.979**	-0.423**	-0.355**	-0.282*
	(0.387)	(0.171)	(0.174)	(0.147)
Observations	738	738	738	738
Elections	82	82	82	82

Note: This table reports the results of the regression discontinuity design. The first two columns show that households of the caste that won the election are less likely to pay any bribes and bribe a smaller fraction of the officials they approach for government services. The third and fourth columns, which explore the reason for the bribe, show that winners are less likely to pay a bribe to get a ration card and the record for rights for their land. Panel A shows the results for the entire sample. Panel B demonstrates that the results hold when I restrict the sample to households used for the manipulation tests in Table 3.4—that is, households for which I have data from 1982. Each observation is a household-election, and all standard errors are clustered at the election.

Table 3.1 give descriptive statistics for several of the variables, each calculated over the largest possible sample. For most of the empirical work I use only elections within ten percentage points of the discontinuity, and for a test of robustness I use elections within five percentage points. The value of crops, shown in 1960 Indian Rupees, converts at current exchange rates to 2157 U.S. dollars.

Many jaded Indians believe all officials are corrupt, but the survey suggests otherwise. For the average household only 11 percent of the officials approached between elections had to be bribed, and only 36 percent of households had to pay a bribe. But the chance of having to pay a bribe is not fixed across all households or all elections. Recall Panel A of Figure 3.1, which shows the probability of having to pay a bribe as a function of the vote difference. The figure shows that losing households are far more likely to pay bribes after close elections.

3.6 Results

Table 3.2 reports the results of using Equation 3.2 to estimate the effect of being in the winning faction on several measures of bribery. The first column of Panel A, which looks at the probability of having to pay any bribes, estimates the size of the gap shown in Figure 3.1. The estimate suggests that the winners of a close election never pay bribes while the losers always pay bribes. The second column shows that winners expect to bribe roughly 38

Table 3.3
Robustness

Polynomial Order	$-10 < VoteDifference < 10$	$-5 < VoteDifference < 5$
Order 2		-1.385***
		(0.351)
Order 3	-0.937***	-1.617***
	(0.336)	(0.323)
Order 4	-1.064**	
	(0.431)	
Order 5	-1.546***	
	(0.301)	
Observations	1054	420
Elections	88	37

Note: I test whether the results are sensitive to how I estimate the regression discontinuity. Each cell reports a different specification's estimate of the effect of being in the winning faction on the probability of having to pay a bribe. Each row varies the order of the polynomial control function and each column the window around the discontinuity.

percentage points fewer of the officials they approach for services. The last two columns look at the probability of paying a bribe for specific services. Winners are 31 and 44 percentage points less likely to pay bribes for ration cards and the Record of Rights to their land.

Why are these estimates so big? My main measure of factionalism, the change in the probability of having to pay any bribes, is as large as it can be. Figure 3.1 suggests that close elections differ from other elections. Away from the discontinuity bureaucrats prey on all factions equally, but near the discontinuity factionalism grows worse. It is beyond the scope of this paper to explore why close elections worsen factionalism. What matters here is that if factionalism distorts production the biggest distortions should be at the discontinuity.

To confirm the results are robust Table 3.3 checks that that the probability of having to pay a bribe remains consistent when I change two parts of the specification. The rows vary the degree of the polynomial used in the control function while the columns vary the window around the discontinuity. Since few elections lie within the smaller window I control for bias there with polynomials of lower order. All of these estimates lie within two standard errors of 1.

Table 3.4 verifies that households barely on either side of the discontinuity do not differ as they would if the rich rigged the election or predicted its outcome. I check whether several pre-determined outcomes jump at the discontinuity. I find no statistically significant change in any of my measures of wealth and success in 1982 or in fixed factors like caste and religion. These tests discard households for whom I have no data from 1982. Could the

Table 3.4
Manipulation Tests: Pre-Determined Variables

	Log Income (1982)	Log Wealth (1982)	Log Savings (1982)	Land (1982)	Age of Head (1982)
Estimate	0.704	0.522	0.900	3.538	-4.811
	(0.611)	(0.871)	(1.170)	(3.140)	(13.379)
Observations	738	738	712	738	738
Elections	82	82	80	82	82

	HH Size (1982)	Scheduled Caste (2006)	High Caste (2006)	Hindu (2006)	Muslim (2006)
Estimate	-1.291	-0.005	-0.277	-0.196	0.110
	(2.883)	(0.111)	(0.809)	(0.135)	(0.070)
Observations	738	1054	1054	1054	1054
Elections	82	88	88	88	88

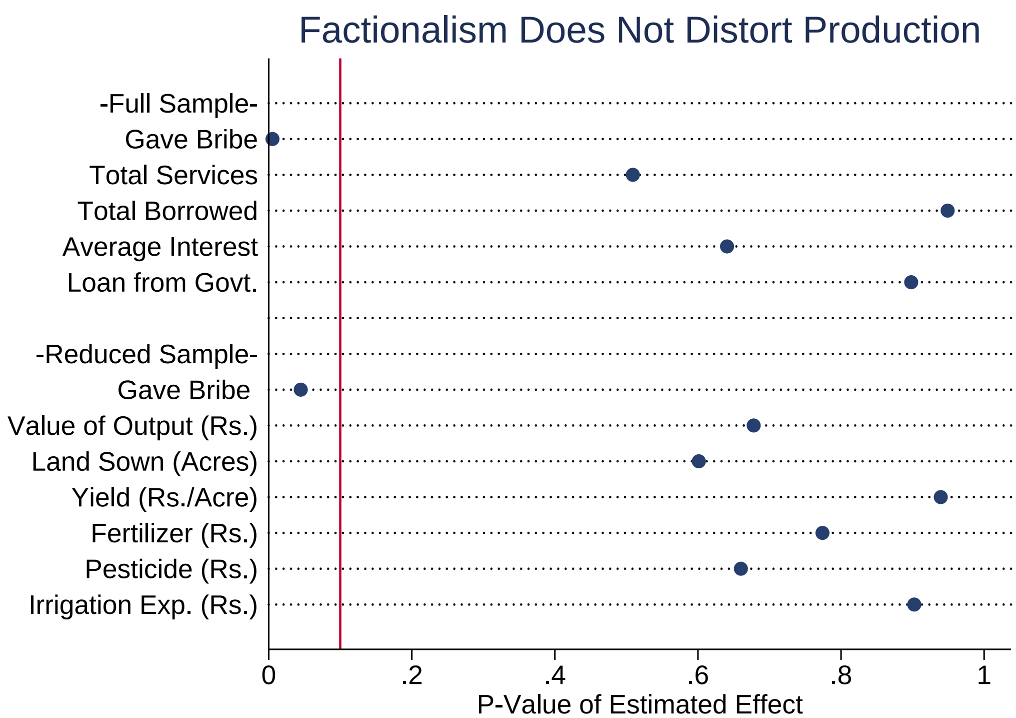
Note: I test whether the outcome of a close election is truly random. Each cell gives the estimate of how some feature of the household changes at the discontinuity. All these features were either recorded in the 1982 round of the survey or unlikely to change after birth (the year the value was recorded is given in parentheses). Households just above the discontinuity do not seem to differ from houesholds just below it.

Table 3.5
Placebo Tests: Fake Discontinuities

	(1) Discontinuity at -5	(2) Discontinuity at 5
[Placebo Disc.]	-4.520	-3.329
	(4.65)	(6.52)
Observations	1054	1054
Elections	88	88

Note: I estimate the effect of “winning” the election under the (false) assumption that a farmer’s candidate wins the election with a vote difference of -5 and 5. The estimated effects at these fake discontinuities are insignificant.

Figure 3.5
P-value on Tests for Whether Output Changes at the Discontinuity



Note: I estimate the regression discontinuity (Equation 3.3) to see if winners produce more output or seek more government services. I test whether each measure changes at the discontinuity—whether the coefficient on $[PastPost]$ is significant—and graph the p-values. The top cluster uses the full sample with 1054 observations and 88 elections. The one exception is the test on average interest, which uses the subset of households that borrowed any money (486 observations and 70 elections). The bottom cluster uses the sample of observations for which I have data on production from 1999 or 2006. This sample contains 566 observations and 54 elections. In both samples I can reject that the probability of paying a bribe does not change but cannot reject that any measure of government services or production does not change.

failure to reject be driven by lost precision? Panel B of Table 3.4 shows that the answer is no; even in the reduced sample I find evidence of factionalism. Table 3.5 tests whether the probability of paying a bribe jumps at two fake discontinuities. As expected neither fake discontinuity causes a statistically significant change.

The important test, however, is whether factionalism distorts production. The theory predicts that neither the take-up of services nor the output of production should change at the discontinuity. The prediction is consistent with Figure 3.5, which shows for each outcome the p-value on the test of a change. The first cluster of results confirms that the probability of paying a bribe does change. But the number of government services sought, the total amount borrowed from all sources, the average interest rate conditional on borrowing, and the probability of getting a loan from a government source do not change. The second cluster of results looks at the restricted sample for which I have data on production.

I confirm that even in the restricted sample the probability of paying a bribe changes. But the value of farm output, the acres of land sown, the yield, and expenditures on fertilizers, pesticides, and irrigation all show no evidence of a significant change. As summarized here and detailed in Table C.3 (see Appendix C.1), the results are consistent with the model.

I find no evidence that factionalism distorts production; this does not prove factionalism cannot distort production. We might worry that my test is too noisy and my standard errors too big to reject a false null. But the noise and error are not too big to the reject that the chance of paying a bribe changes at the discontinuity. The estimates from Table 3.2 show a big difference in how bureaucrats treat winners and losers when asking for bribes. If factionalism distorted production we would expect an equally big difference in output. My results show that even when factionalism is as bad as it can be—after close elections, when bureaucrats never ask winners but always ask losers for bribes—any distortion caused is too small to detect.

The larger question is whether factionalism ever distorts production. Is there anything about an Indian village that makes the assumptions of Section 3.2 uniquely valid? The first assumption is that the service provided is either non-rival or not scarce. The government can give every farmer the Record of Rights to her own land and a connection to the electricity grid. The government cannot give every firm a the rights to mine a public resource. A favored firm will get these rights even if it cannot put them to best use.

The second assumption is that bureaucrats know the value of their services—not just the average value but the exact value for each firm. Knowing the firm's willingness to pay lets the bureaucrat set a bribe that does not distort the firm's actions. Even if the firm has a credit constraint the model's prediction holds as long as the bureaucrat knows the constraint. He sets the highest bribe the firm is both willing and able to pay. The assumption of perfect information suits a small Indian village. The land secretary knows how much land each farmer owns; he knows how large a loan she can get if she proves her land is valid collateral. He can guess how much she is willing to pay to get the Record of Rights. Perfect information is less likely in national government or when the bureaucrat sells services needed by firms that produce output that is unfamiliar. That said, letting the firm bargain with the bureaucrat might obviate the assumption if bargaining gives the bureaucrat information. My results suggest future research should test whether better informed bureaucrats demand more efficient bribes.

3.7 Conclusion

Though distasteful, factionalism need not distort production. I build a model in which bureaucrats can demand bribes for a government service that makes firms more productive. Bureaucrats are less likely to be caught demanding bribes from a firm in the losing faction. Nevertheless they choose bribes that do not distort production from losers to winners. I test the model in rural Indian villages using a regression discontinuity design. I show that households of the same caste as the village council president pay fewer bribes for government services. But as the model predicts, I find no evidence to suggest they produce more output.

My results differ from earlier work because the setting behind both my theory and empirics differ. Unlike the citizens in Banerjee (1997), farmers in India need services that are non-rival. Unlike the government in Acemoglu and Verdier (2000), the government in an Indian village does not want to correct a market failure. And unlike the bureaucrats in either model, the bureaucrats in rural India know the value of their services. Future research must test whether changing these conditions changes the costs of factionalism.

APPENDIX A

Appendix: Market Failures and Misallocation

A.1 Proof of the Bounding Condition

I prove that my measure of factor market misallocation is a lower bound on the truth. Call the land-capital and labor-capital ratios derived with the endowment assumption $\tau = \frac{T_i}{K_i}, \Lambda = \frac{L_i}{K_i}$, and define the ratios without the assumption $\tilde{\tau}, \tilde{\Lambda}$ similarly. According to (1.2) they must be identical for all farmers. Consider the market-clearing condition for land with the endowment assumption:

$$\begin{aligned}
\sum T_i^+ &= T_I \\
\Rightarrow \sum K_i^+ \tau &= T_I \\
\Rightarrow \tau K_I &= T_I \\
\Rightarrow \tau &= T_I / K_I
\end{aligned}$$

Identical reasoning shows $\tilde{\tau} = T_I / K_I$ as well, so $\tau = \tilde{\tau}$ and similarly $\Lambda = \tilde{\Lambda}$

The difference between aggregate output with and without the endowment condition is

$$
\begin{aligned}
\tilde{Y}_I^+ - Y_I^+ &= \sum A_i\phi_i(K_i^+)^{\theta_K}(T_i^+)^{\theta_T}(L_i^+)^{\theta_L} - \sum A_i\phi_i(\tilde{K}_i^+)^{\theta_K}(\tilde{T}_i^+)^{\theta_T}(\tilde{L}_i^+)^{\theta_L} \\
&= \sum A_i\phi_i(K_i^+)^{\sigma}\tau^{\theta_T}\Lambda^{\theta_L} - \sum A_i\phi_i(\tilde{K}_i^+)^{\sigma}\tau^{\theta_T}\Lambda^{\theta_L} \\
&= \tau^{\theta_T}\Lambda^{\theta_L}\sum A_i\phi_i\left[(K_i^+)^{\sigma} - \tilde{K}_i^+)^{\sigma}\right] \\
\Rightarrow \frac{1}{Y_I}\left[(\tilde{Y}_I^+ - Y_I) - (Y_I^+ - Y_I)\right] &= \frac{1}{Y_I}\tau^{\theta_T}\Lambda^{\theta_L}\sum A_i\phi_i\left[(K_i^+)^{\sigma} - \tilde{K}_i^+)^{\sigma}\right] \\
\Rightarrow \tilde{G}_I^{FACT} - G_I^{FACT} &= \frac{\tau^{\theta_T}\Lambda^{\theta_L}}{Y_I}\sum A_i\phi_i\left[(K_i^+)^{\sigma} - \tilde{K}_i^+)^{\sigma}\right] \\
\Rightarrow \mathbb{E}_i[\mathbb{E}_\phi[\tilde{G}_I^{FACT} - G_I^{FACT}]] &= \frac{\tau^{\theta_T}\Lambda^{\theta_L}|I|}{Y_I}\mathbb{E}_i\left(A_i\left[(K_i^+)^{\sigma} - \tilde{K}_i^+)^{\sigma}\right]\right)
\end{aligned}
$$

where the last step follows because unanticipated productivity ϕ is independent and mean 1, and the size of village I is $|I|$. Since $\frac{\tau^{\theta_T}\Lambda^{\theta_L}|I|}{Y_I} > 0$, $\mathbb{E}_i[\mathbb{E}_\phi[\tilde{G}_I^{FACT} - G_I^{FACT}]] > 0$ if and only if $\mathbb{E}_i\left(A_i\left[(K_i^+)^{\sigma} - \tilde{K}_i^+)^{\sigma}\right]\right) > 0$. In words, G_I^{FACT} underestimates the gains from perfect factor markets. Since $G_I = G_I^{FACT} + G_I^{FIN}$ it must be that G_I^{FIN} overestimates the subsequent gains from financial market perfection. ■

A.2 Robustness: Subdistrict-Level Reallocation

This appendix explains how to calculate misallocation at the sub-district level. The firm's problem is the same as before, and its optimal capital choice is $K_i^* = \eta A_i^{\frac{1}{1-\sigma}}$ where $\eta = \left[\left(\frac{\theta_K}{w^K}\right)^{1-\theta_T-\theta_L}\left(\frac{\theta_T}{w^T}\right)^{\theta_T}\left(\frac{\theta_L}{w^L}\right)^{\theta_L}\right]^{\frac{1}{1-\sigma}}$. Replace the village-level market-clearing condition with village and subdistrict-level conditions:

$$
\begin{aligned}
\sum_{i\in I} K_i^* = K_I^* \qquad \forall I \\
\sum_I w_I K_I^* = \sum_I w_I \bar{K}_I
\end{aligned}
$$

where K_I^* is the amount of capital the village is optimally allocated, $\bar{K}_I$ is the village's initial allocation, and w_I is an inverse-probability weight (total number of households in the village divided by the number of households sampled). Then

$$
\begin{aligned}
\eta \sum_I w_I \sum_{i \in I} A_i^{\frac{1}{1-\sigma}} &= \sum_I w_I \bar{K}_I \\
\Rightarrow \eta &= \frac{\sum_I w_I \sum_I w_I \bar{K}_I}{\sum_I w_I \sum_{i \in I} A_i^{\frac{1}{1-\sigma}}}
\end{aligned}
$$

Sub this back into the individual demand:

$$
K_i^* = \frac{A_i^{\frac{1}{1-\sigma}}}{\sum_I w_I \sum_{i \in I} A_i^{\frac{1}{1-\sigma}}} \sum_I w_I \bar{K}_I
$$

Let y_i^* be output with the optimal capital, land, and labor. Optimal aggregate output is

$$
Y^* = \sum_I w_I \sum_{i \in I} y_i^*
$$

A.3 Technical Appendix: General Model

This appendix presents a more general form of the model I use in Section 1.3 to derive my measures of misallocation, and also derives their asymptotic behavior. Suppose households are the sole economic actors, and as ever they live to maximize their lifetime utility from consumption. They earn income by selling or renting out the factors they own (including labor) and by operating firms. Aside from the usual budget constraint, they face potentially binding constraints on their choices of factors. For example, if no labor market exists they are constrained to use exactly their labor endowment. They may also be constrained in their period-to-period liquidity. They save and borrow at interest rates that need not be common across households, and may also have to pay an external finance premium to borrow. A household's access to insurance may be imperfect, which means its consumption depends on the profits of its firm. Finally, households differ in their preferences (notably their risk tolerance) and the productivity of the firms they operate.

Suppose household i owns and operates firm $i \in I_t$, where I_t is some group of firms (a village or a sector). The household maximizes present discounted utility from consumption over an infinite horizon:

$$\max_{(c_{i,t+j},\boldsymbol{X}_{i,t+j},\boldsymbol{I}_{i,t+j})_{j=0}^{\infty}} \mathbb{E}[\sum_{j=0}^{\infty} \rho^j u_i(c_{i,t+j}) \mid \mathcal{I}_{it}].$$

Subject to:

$$\begin{aligned}
c_{it} + b_{i,t+1} &= y_{it} + [1 + r_{it} + \zeta_{it}(z_{it} - b_{it}; b_{it}, \boldsymbol{X}^o_{i,t-1})](z_{it} - b_{it}) && (\mathit{Budget\ Const.})\\
y_{it} &= f(\boldsymbol{A}_{it}, \boldsymbol{\phi}_{it}, \boldsymbol{X}_{it}; \boldsymbol{\theta}_i) && (\mathit{Production})\\
z_{it} &= \boldsymbol{w}_{it}^T(\mathbf{X}_{it} - \boldsymbol{X}^o_{it}) + \boldsymbol{p}_{it}^T \boldsymbol{I}_{it} && (\mathit{Expenditure})\\
X^o_{k,it} &= (1 - \delta^k) X^o_{k,i,t-1} + I_{k,it} && \forall k = 1, \ldots, K\\
z_{it} - b_{it} &\leq \omega_{it}(b_{it}, \boldsymbol{X}^o_{i,t-1}) && (\mathit{Liquidity\ Const.})\\
\underline{\mathbb{X}}_{it} &\leq \mathbb{X}_{it}(\boldsymbol{X}_{it} - \boldsymbol{X}^o_{it}, \boldsymbol{I}_{it}) \leq \overline{\mathbb{X}}_{it} && (\mathit{Factor\ Choice\ Const.})
\end{aligned}$$

where $\mathcal{I}$ is the information set, f a strictly concave decreasing returns revenue production function, $\boldsymbol{A}$ anticipated revenue productivity, $\boldsymbol{\phi}$ unanticipated revenue productivity, $\boldsymbol{\theta}$ a vector of production parameters, $\boldsymbol{X}$ factor levels used in production, c is consumption, b borrowing, r the borrowing rate, $\zeta(\cdot)$ an external finance premium, y_{it} revenue, $\boldsymbol{X}^o$ owned factors, $\boldsymbol{I}$ purchase of factors, $\boldsymbol{p}$ a vector of factor purchase prices, z input expenditure, and $\omega(\cdot)$ a liquidity constraint. $\mathbb{X}_{it}$ is a continuously twice-differentiable factor choice transformation function, and $\underline{\mathbb{X}}_{it}, \overline{\mathbb{X}}_{it}$ upper and lower bounds on (transformed) factor choice;

they bound a household's access to factors beyond those it owns (rented factors). Assume all past and currently dated variables are elements of $\mathcal{I}$ except ϕ. Rental prices $\boldsymbol{w}$ and all other prices can vary by household/firm i.

For notational simplicity I model insurance markets implicitly as the correlation between a household's unanticipated productivity and its consumption (perfect insurance ensures zero correlation). I have assumed away output and asset taxes because they are not important in the empirical application; accounting for them is straightforward if the tax schedule is known.[1]

Let X_{k,I_t} be the aggregate stock of factor k among the unit measure of firms in I_t. Define an *allocation vector* as a set of K-dimensional vectors $\{\mathbf{X}'_{it}\}_{i \in I_t}$ such that $\int_{i \in I_t} X'_{k,it}\, di = X_{k,I_t} \forall k$.

Varying factor prices and savings rates, external finance premiums, liquidity constraints, and factor choice constraints can all distort realized allocations away from the frictionless benchmark. Eliminating them separates the household problem from the firm problem and produces the production allocations of the frictionless neoclassical world. Denote outcomes in the world with no constraints or market imperfections by asterisks, and characterize it with these conditions:

$$\begin{aligned}
\text{(Law of One Price) } & \boldsymbol{w}_{it} = \boldsymbol{w}_{I_t}, \boldsymbol{p}_{it} = \boldsymbol{p}_{I_t} \forall i \in I_t m_k \forall k \\
\text{(Unconstrained Factor Choices) } & \underline{\mathbb{X}}_{it} = \mathbb{X}(-\boldsymbol{X}^o_{it}, -\boldsymbol{X}^o_{i,t-1}), \overline{\mathbb{X}}_{it} = \mathbb{X}(\infty, \infty) \ \ \forall i \in I_t \\
\text{(Perfect Credit Markets) } & \omega_{it} = \infty, \zeta_{it}(\cdot) = 0, r_{it} = r_{I_t} \ \ \forall i \in I_t \\
\text{(Perfect Insurance Markets) } & c_{it} \perp \boldsymbol{\phi}_{it} \ \ \forall i \in I_t.
\end{aligned}$$

Under these assumptions the firm maximizes per-period expected profit independently of the household's dynamic consumption problem:

$$\max_{\boldsymbol{X}_{it}} \qquad \mathbb{E}[f(\boldsymbol{A}_{it}, \boldsymbol{\phi}_{it}, \boldsymbol{X}_{it}; \boldsymbol{\theta}_i) - \boldsymbol{w}^T_{I_t} \boldsymbol{X}_{it} \quad | \mathcal{I}_{it}]$$

Then the following first-order conditions and market-clearing conditions characterize

[1] For output taxes, for example, one would simply modify the budget constraint to be $c_{it} + b_{i,t+1} = (1 - \tau_{it}) y_{it} + \cdots$ and then perform all subsequent operations conditional on the presence of the taxes to account for the fact that they will continue to distort even the counterfactual optimal scenarios where market failures are eliminated.

the unique general equilibrium outcome:

$$\mathbb{E}[f_{X_{k,it}}(\boldsymbol{A}_{it}, \boldsymbol{\phi}_{it}, \boldsymbol{X}^*_{it}; \boldsymbol{\theta}_i)] = w_{k,I_t} \quad \forall k \tag{A.1}$$

$$\int_{i \in I_t} X^*_{k,it}(w_{k,I_t}, \boldsymbol{A}_{it}, X^*_{-k,it}; \boldsymbol{\theta}_i)\, di = X_{k,I_t} \quad \forall k \tag{A.2}$$

Solving these equations solves for the optimal allocation vector $\{\boldsymbol{X}^*_{it}\}$. The optimal allocations solve a system of equations that contain only observables and production parameters estimable from observables. This makes calculating the counterfactual scenario with production and factor data possible.

To solve for the outcome where factor markets are perfect I must assume unanticipated shocks are Hicks-Neutral. That is, $y_{it} = \phi_{it} f(\boldsymbol{A}_{it}, \boldsymbol{X}_{it}; \boldsymbol{\theta}_i)$. Since all factors are equally risky in production, imperfect insurance only affects overall expenditure and not expenditure on capital versus labor. Consider the following hypothetical: Firm $i \in I_t$, which uses $\bar{\boldsymbol{X}}_{it}$ in production, now has those factors as "endowments." Each firm can then trade factors with $-i \in I$ subject to its expenditures being equal to the value of its endowment until its factor mix is optimal.

Since firms cannot change their total expenditure for a period, they again optimize period-by-period:

$$\max_{\boldsymbol{X}_{it}} \qquad \mathbb{E}[\phi_{it} f(\boldsymbol{A}_{it}, \boldsymbol{X}_{it}; \boldsymbol{\theta}_i) - \boldsymbol{w}^T_{I_t} \boldsymbol{X}_{it}]$$

Subject to:

$$\boldsymbol{w}^T_{I_t}(\boldsymbol{X}_{it} - \bar{\boldsymbol{X}}_{it}) = 0$$

Essentially, I have dropped the firms into an Edgeworth economy where their profit function plays the role of a utility function. The following equations characterize the unique outcome:

$$\frac{\mathbb{E}[f_{X_k}(\boldsymbol{A}_{it}, \boldsymbol{X}^+_{it}; \boldsymbol{\theta}_i)]}{\mathbb{E}[f_{X_j}(\boldsymbol{A}_{it}, \boldsymbol{X}^+_{it}; \boldsymbol{\theta}_i)]} = \frac{w_{k,I_t}}{w_{j,I_t}} \quad \forall k, j\ \forall i \tag{A.3}$$

$$\boldsymbol{w}^T_{I_t}(\boldsymbol{X}_{it} - \bar{\boldsymbol{X}}_{it}) = 0 \quad \forall i \tag{A.4}$$

$$\int_{i \in I_t} X^+_{k,it}(w_{k,I_t}, \boldsymbol{A}_{it}, X^+_{-k,it}; \boldsymbol{\theta}_i)\, di = X_{k,I_t} \quad \forall k \tag{A.5}$$

Define efficiency and the gains from reallocation as in the main text. Assume anticipated productivity is also Hicks-Neutral, the production function is homogeneous, and common production parameters. I can prove this theorem about my measure of the costs of factor versus financial market failures:

Proposition 2 (Bounding) *Assume $y_{it} = A_{it}\phi_{it} f(\boldsymbol{X}_{it}; \boldsymbol{\theta}_i)$, f is homogeneous of degree σ, $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{I_t}\ \ \forall i \in I_t$, and $\mathbb{E}[A_{it}\{(\tilde{X}^+_{1,it})^\sigma - (X^+_{1,it})^\sigma\}] > 0$, where $\tilde{X}_{1,it}$ is the quantity of the first factor i would choose at time t under Unconstrained Factor Choices and the Law of One Price without being constrained by endowments. Then G^{FACT}_I is a lower-bound on the true gains from moving to Unconstrained Factor Choices and the Law of One Price. Likewise, G^{FIN}_I is an upper-bound on the true gains from subsequently creating Perfect Credit and Insurance Markets.*

Proof: Consider the (unobservable and incalculable) outcome where the Law of One Price and Unconstrained Factor Choices hold without the extra restriction of endowment conditionality. Call this the True Perfect Factor Market outcome. Denote with superscript $+$ a variable specific with perfect factor markets and the endowment condition: the Endowment-Conditional Perfect Factor Market (ECPFM) outcome. Let overset tilde and superscript $+$ variables come from the True Perfect Factor Market (TPFM) outcome. An over-bar denotes variables from to the observed/realized outcome. For notational ease, suppress the common production parameter $\boldsymbol{\theta}_I$. I first prove two lemmas useful to the main result.

Lemma 1 *For any level or vector of factor choices X_{it}, let $\ddot{X}_{it} = X_{it}/X_{1,it}$. Then $\ddot{\tilde{\boldsymbol{X}}}^+_{it} = \ddot{\tilde{\boldsymbol{X}}}^+_{I_t}$ and $\ddot{\boldsymbol{X}}^+_{it} = \ddot{\boldsymbol{X}}^+_{I_t}$ for all $i \in I$ (that is, all firms in both outcomes will employ factors in exactly the same proportions).*

The optimality condition for ECPFM is

$$\frac{f_{X_k}(\boldsymbol{X}^+_{it})}{f_{X_1}(\boldsymbol{X}^+_{it})} = \frac{w^+_k}{w^+_1}.$$

By homogeneity,

$$\begin{aligned}\frac{X^{\sigma-1}_{1,it} f_{X_k}(\boldsymbol{X}^+_{it}/X_{1,it})}{X^{\sigma-1}_{1,it} f_{X_1}(\boldsymbol{X}^+_{it}/X_{1,it})} &= \frac{w^+_k}{w^+_1}\ \ \forall k \\ \Rightarrow \frac{f_{X_k}(\ddot{\boldsymbol{X}}^+_{it})}{f_{X_1}(\ddot{\boldsymbol{X}}^+_{it})} &= \frac{w^+_k}{w^+_1}\ \ \forall k\end{aligned}$$

Since f satisfies strictly decreasing returns, $\boldsymbol{X}_{it}$ is unique and thus the $\ddot{\boldsymbol{X}}_{it}$ that satisfies the above conditions is also unique for each i, t. But the above conditions are not functions of any variables unique to i (e.g. A_{it}), and thus $\ddot{\boldsymbol{X}}_{it} = \ddot{\boldsymbol{X}}_{I_t}$ for all $i \in I_t$. A similar argument shows that $\ddot{\tilde{\boldsymbol{X}}}_{it} = \ddot{\tilde{\boldsymbol{X}}}_{I_t}$ for all $i \in I_t$. ■

Lemma 2 $\ddot{\tilde{\boldsymbol{X}}}^{+}_{I_t} = \ddot{\boldsymbol{X}}^{+}_{I_t}$ *for all* $i \in I_t$*, that is the mixes of factors will be identical with TPFM and ECPFM.*

Since $\{\boldsymbol{X}^{+}_{it}\}$ and $\{\tilde{\boldsymbol{X}}^{+}_{it}\}$ are both allocation vectors, each factor must aggregate to the total observed stock. For the latter, for example, for all k

$$
\begin{aligned}
& & \int_{i \in I_t} X^{+}_{k,it} di = X_{k,It} \\
\Rightarrow & & \int_{i \in I_t} X_{1,it} \ddot{X}^{+}_{k,I_t} di = X_{k,I_t} \\
\Rightarrow & & \ddot{X}^{+}_{k,I_t} \int_{i \in I_t} X_{1,it} di = X_{k,I_t} \\
\Rightarrow & & \ddot{X}^{+}_{k,I_t} X_{1,It} = X_{k,I_t} \\
\Rightarrow & & \ddot{X}^{+}_{k,I_t} = \frac{X_{k,I_t}}{X_{1,I_t}}
\end{aligned}
$$

where the second line follows from Lemma 1. Parallel arguments show that $\ddot{\tilde{X}}^{+}_{k,It} = \frac{X_{k,It}}{X_{1,It}}$. Then

$$
\ddot{X}^{+}_{k,It} = \ddot{\tilde{X}}^{+}_{k,It} \quad \forall k
$$

implying $\ddot{\boldsymbol{X}}^{+}_{It} = \ddot{\tilde{\boldsymbol{X}}}^{+}_{It}$. ■

To prove the main result, write the difference between the TPFM and ECPFM outcome as

$$
\begin{aligned}
\tilde{Y}^{+}_{I_t} - Y^{+}_{I_t} &= \int_{i \in I_t} A_{it} \phi_{it} f(\tilde{\boldsymbol{X}}^{+}_{it})\, di - \int_{i \in I_t} A_{it} \phi_{it} f(\boldsymbol{X}^{+}_{it})\, di \\
&= \mathbb{E}[A_{it} \phi_{it} f(\tilde{\boldsymbol{X}}^{+}_{it})] - \mathbb{E}[A_{it} \phi_{it} f(\boldsymbol{X}^{+}_{it})] \\
&= \mathbb{E}[A_{it} \phi_{it} \{f(\tilde{\boldsymbol{X}}^{+}_{it}) - f(\boldsymbol{X}^{+}_{it})\}] \\
&= \mathbb{E}[\phi_{it}] \mathbb{E}[A_{it} \{f(\tilde{\boldsymbol{X}}^{+}_{it}) - f(\boldsymbol{X}^{+}_{it})\}] \\
&= \mathbb{E}[A_{it} \{f(\tilde{\boldsymbol{X}}^{+}_{it}) - f(\boldsymbol{X}^{+}_{it})\}]
\end{aligned}
$$

where the second equality comes from the measure 1 normalization, the fourth the independence of anticipated and unanticipated variables, and the fifth the unit mean normalization of ϕ_{it}.

By the homogeneity of f and Lemma 1,

$$f(\tilde{\boldsymbol{X}}_{it}^{+}) = (\tilde{X}_{1,it}^{+})^{\sigma} f(\ddot{\tilde{\boldsymbol{X}}}_{I_t}^{+})$$
$$f(\boldsymbol{X}_{it}^{+}) = (X_{1,it}^{+})^{\sigma} f(\ddot{\boldsymbol{X}}_{I_t t}^{+})$$

and by Lemma 2 $f(\ddot{\boldsymbol{X}}_{it}^{+}) = f(\ddot{\tilde{\boldsymbol{X}}}_{it}^{+}) = a$. Applying these results to the above expressions, we have that

$$\begin{aligned}
\tilde{Y}_{I_t}^{+} - Y_{I_t}^{+} &= \mathbb{E}[A_{it}\{f(\tilde{\boldsymbol{X}}_{it}^{+}) - f(\boldsymbol{X}_{it}^{+})\}] \\
&= \mathbb{E}[A_{it}\{(\tilde{X}_{1,it}^{+})^{\sigma} a - (X_{1,it}^{+})^{\sigma} a\}] \\
&= a\mathbb{E}[A_{it}\{(\tilde{X}_{1,it}^{+})^{\sigma} - (X_{1,it}^{+})^{\sigma}\}]
\end{aligned}$$

Since $a > 0$ and $\mathbb{E}[A_{it}\{(\tilde{X}_{1,it}^{+})^{\sigma} - (X_{1,it}^{+})^{\sigma}\}] > 0$ by assumption, TPFM output is greater than in the ECPFM outcome, and so the calculated gains will be as well.

QED

Now suppose that consistent estimators of $\{\boldsymbol{A}_{it}, \boldsymbol{\phi}_{it}, \boldsymbol{\theta}_i\}$ are available. Then it is a numerical exercise to solve the sample analogs of (A.1) and (A.2) for estimates of the CCM allocations and (A.3), (A.4), and (A.5) for sample analogs of the PFM allocations. Plug the computed allocations into the expressions for efficiency and the gains. The following proposition summarizes the asymptotic properties of these estimators:

Proposition 3 *Suppose $\hat{I}$ is a random sample of I, and define $\hat{E}_{\hat{I}}, \hat{E}_{\hat{I}}^{FACT}, \hat{E}_{\hat{I}}^{FIN}$, and the estimators of the other components as described above. Finally, assume the expectations and variances of $\boldsymbol{A}_{it}, \boldsymbol{\phi}_{it}, \boldsymbol{\theta}_i, y_{it}, \boldsymbol{X}_{it}$ are finite. Then the estimators are all consistent and asymptotically normal.*

Proof:

Consistency: I will prove the consistency of $\hat{E}_{\hat{I}_t}$; demonstrating the consistency of the other estimators is similar. Suppress time subscripts for notational simplicity. I will first identify the population parameters in terms of their moments, and then demonstrate that the sample analogs are consistent estimators.

Recall that (A.1) characterizes any interior solution to the population optimization - in other words, if an interior solution exists, the function $\boldsymbol{X}^{*}(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \boldsymbol{w}_I)$ characterizes firm i's optimal allocation as a function of i-specific parameters and the prices. Since f is concave and satisfies DRS, the solution is not only interior but also unique. Moreover, the Maximum Theorem Under Convexity (see Sundaram, 1996, p. 237) guarantees that $\boldsymbol{X}^{*}$

is a continuous function of the population prices $\boldsymbol{w}_I$ (see the Lemma below). Inspection demonstrates that the optimality conditions are identical for the sample optimization, and thus the derived $\boldsymbol{X}^*$ is as well.

Applying the measure-one normalization, (A.2) in population reduces to

$$\mathbb{E}[X^*(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \boldsymbol{w}_I)] = \mathbb{E}[X_i].$$

Note that the sample analog of (A.2) in random sample $\hat{I}$ reduces to the sample analog of this moment condition trivially:

$$\begin{aligned} \sum_{i \in \hat{I}} X^*(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{w}_I) &= \sum_{i \in \hat{I}} \boldsymbol{X}_i \\ \frac{\sum_{i \in \hat{I}} X^*(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\theta}}_i, \boldsymbol{w}_I)}{|\hat{I}|} &= \frac{\sum_{i \in \hat{I}} \boldsymbol{X}_i}{|\hat{I}|} \end{aligned}$$

Since $\{\hat{\boldsymbol{A}}_{it}\}, \{\hat{\boldsymbol{\theta}}_i\}$ are consistent estimators for the population technology parameters, then together with the Lemma below this implies that $\hat{\boldsymbol{w}}_I$ is a consistent GMM estimator for $\boldsymbol{w}_I$.

Applying the measure 1 normalization to the definition of E, we have

$$E_I = \frac{\mathbb{E}[y_i]}{\mathbb{E}[f(\boldsymbol{A}_i, \boldsymbol{\phi}_i, \boldsymbol{X}^*(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \boldsymbol{w}_I); \boldsymbol{\theta}_i)]}$$

while the estimator of E is

$$\begin{aligned} \hat{E}_{\hat{I}} &= \frac{\sum_{i \in \hat{I}} y_i}{\sum_{i \in \hat{I}} f(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\phi}}_i, \boldsymbol{X}^*(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{w}}_I); \hat{\boldsymbol{\theta}}_i)} \\ &= \frac{\sum_{i \in \hat{I}} y_i}{|\hat{I}|} \Big/ \frac{\sum_{i \in \hat{I}} f(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\phi}}_i, \boldsymbol{X}^*(\hat{\boldsymbol{A}}_i, \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{w}}_I); \hat{\boldsymbol{\theta}}_i)}{|\hat{I}|}. \end{aligned}$$

Since $\hat{I}$ is a random sample and y_i has finite expectation, the numerator of $\hat{E}$ is consistent for $\mathbb{E}[y_i]$ by Kolmogorov's Law of Large Numbers. And since f and $\boldsymbol{X}^*$ are continuous in all their arguments and $\{\hat{\boldsymbol{A}}_{it}\}, \{\hat{\boldsymbol{\phi}}_{it}\}, \{\hat{\boldsymbol{\theta}}_i\}, \hat{\boldsymbol{w}}_{\hat{I}}$ are all consistent for their respective population parameters, the denominator is consistent for $\mathbb{E}[f(\boldsymbol{A}_i, \boldsymbol{\phi}_i, \boldsymbol{X}^*(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \boldsymbol{w}_I); \boldsymbol{\theta}_i)]$ by Kolmogorov's Law of Large Numbers and the Continuous Mapping Theorem. Then applying the Continuous Mapping Theorem again, the ratio of the two consistent estimators is consistent for the population ratio. Thus, $\hat{E}_{\hat{I}}$ is consistent for E_I.

Lemma (Price Estimator and GMM Consistency): Consider each of the requirements for consistency in turn.

Consistent Weighting Matrix: There are exactly as many market-clearing conditions as prices, implying the estimator is just-identified and thus the weighting matrix irrelevant.

Global Identification: Recall that the optimal outcome will be identical to the solution to the planner's problem. By the assumption that f satisfies decreasing returns, this will be a strictly concave optimization with a unique maximizer $\{\boldsymbol{X}_i\}$. By (A.1), the condition $f_{\boldsymbol{X}}(\boldsymbol{X}_i^*) = \boldsymbol{w}$ is satisfied for all i. Since $f_{\boldsymbol{X}}$ is a function, the uniqueness of $\boldsymbol{X}_i^*$ implies the uniqueness of $\boldsymbol{w}$. Thus, $\boldsymbol{w}$ uniquely satisfies the market-clearing conditions.

$\boldsymbol{X}^*(\boldsymbol{w})$ Is Continuous at all $\boldsymbol{w}$: Observe that $\boldsymbol{X}^*$ is the solution to i's optimization problem, and thus it suffices to show the conditions of the Maximum Theorem hold (see Sundaram, 1996, p. 237). Observe that since f is assumed continuous in $\boldsymbol{X}_i$, the continuity condition is satisfied, so one need only show that the constraint set is a compact-valued continuous function of $\boldsymbol{w}$. The firm is implicitly constrained to choose positive values of all factors, so **0** is a lower bound. Meanwhile, since f satisfies decreasing returns, for each w_k there exists some $\tilde{X}_k(w_k)$ such that $f(0, \ldots, \tilde{X}_k(w_k), \ldots, 0) - w_k \tilde{X}_k(w_k) = -100$. Define $\mathcal{W}(\boldsymbol{w}) = \sum_k w_k \tilde{X}_k(w_k)$. Since the firm always can choose zero of all factors and earn profit zero, we can impose that $\boldsymbol{w}^T \boldsymbol{X}_i \leq \mathcal{W}(\boldsymbol{w})$ and the outcome will be identical to that of the unconstrained problem. This "budget constraint" is like any other from consumer theory and thus continuous, and is closed and bounded (thus compact) for all $\boldsymbol{w}$. Thus, $\boldsymbol{X}^*$ is continuous by the Maximum Theorem.

$\boldsymbol{w} \in \Theta$,Which Is Compact: Since $f_{\boldsymbol{X}} > 0$ by assumption, $\boldsymbol{w} >> 0$. Then some $\varepsilon > 0$ exists such that $\boldsymbol{w} >> (\varepsilon, \ldots, \varepsilon)$. Meanwhile, aggregate demand $\mathbb{E}[X_k^*(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \boldsymbol{w}_I)]$ for any factor k is continuous and strictly decreasing in w_k. Then some $\tilde{\boldsymbol{w}}$ exists such that $\mathbb{E}[X^*(\boldsymbol{A}_i, \boldsymbol{\theta}_i, \tilde{\boldsymbol{w}}_I)] = \mathbb{E}[X_i]/2$, and $\tilde{w}_k > w_k$ for all k. Then $\boldsymbol{w} \in [\varepsilon, w_1] \times \cdots \times [\varepsilon, w_K]$, a closed and bounded subset of $\mathbb{R}^K$, which is thus compact.

$\mathbb{E}[\sup_{\boldsymbol{w} \in \Theta} ||\boldsymbol{X}^*(\boldsymbol{w})||] < \infty$: Note that $\boldsymbol{X}^*$ is determined by the satisfaction of (A.1) (and the non-negativity constraint). Since f_{X_k} is strictly decreasing (by strict concavity) and strictly positive (by assumption), for any finite $\boldsymbol{w}$, either the condition will be satisfied by some finite positive $\boldsymbol{X}^*$ or the non-negativity constraint will bind and $\boldsymbol{X}^*$ will have one or more zero elements. Observe that Θ as defined above is closed and bounded, so any $\boldsymbol{w} \in \Theta$ is finite.

Asymptotic Normality: I again prove the result only for $\hat{E}_{\hat{I}}$; similar algebra and applications of limiting statistics prove the result for the other estimators.[2] Suppress time

[2]For example, the gains from reallocation are actually a continuous function of E: $\hat{\mathbb{G}}_{\hat{I}_t} = \frac{1}{\hat{E}_{\hat{I}}} - 1$. Simply

subscripts for notational simplicity.

$$\begin{aligned}
\sqrt{|\hat{I}|}(\hat{E}_{\hat{I}} - E_I) &= \sqrt{|\hat{I}|}\left(\frac{\sum_{i\in\hat{I}} y_i}{\sum_{i\in\hat{I}} y_i^*} - \frac{\mathbb{E}[y_i]}{\mathbb{E}[y_i^*]}\right) \\
&= \sqrt{|\hat{I}|}\left(\frac{\mathbb{E}[y_i^*]\sum_{i\in\hat{I}} y_i - \mathbb{E}[y_i]\sum_{i\in\hat{I}} y_i^*}{\mathbb{E}[y_i^*]\sum_{i\in\hat{I}} y_i^*}\right) \\
&= \sqrt{|\hat{I}|}\left(\frac{\mathbb{E}[y_i^*]\bar{y}_i - \mathbb{E}[y_i]\bar{y}_i^*}{\mathbb{E}[y_i^*]\bar{y}_i^*}\right) \\
&= \sqrt{|\hat{I}|}(\mathbb{E}[y_i^*]\bar{y}_i - \mathbb{E}[y_i^*]\mathbb{E}[y_i] + \mathbb{E}[y_i^*]\mathbb{E}[y_i] - \mathbb{E}[y_i]\bar{y}_i^*)(\mathbb{E}[y_i^*]\bar{y}_i^*)^{-1} \\
&= \sqrt{|\hat{I}|}\Big[\mathbb{E}[y_i^*](\bar{y}_i - \mathbb{E}[y_i]) - \mathbb{E}[y_i](\bar{y}_i^* - \mathbb{E}[y_i^*])\Big](\mathbb{E}[y_i^*]\bar{y}_i^*)^{-1} \\
&= \Big[\mathbb{E}[y_i^*]\cdot\underbrace{\sqrt{|\hat{I}|}(\bar{y}_i - \mathbb{E}[y_i])}_{\mathbb{A}} - \mathbb{E}[y_i]\cdot\underbrace{\sqrt{|\hat{I}|}(\bar{y}_i^* - \mathbb{E}[y_i^*])}_{\mathbb{B}}\Big]\underbrace{(\mathbb{E}[y_i^*]\bar{y}_i^*)^{-1}}_{\mathbb{C}}
\end{aligned}$$

By Kolmogorov's Law of Large Numbers and the Continuous mapping theorem, $\mathbb{C} \xrightarrow{p} \mathbb{E}[y_i^*]^{-2}$. By the Lindeberg-Lévy Central Limit Theorem, $\mathbb{A} \xrightarrow{d} N(0, \sigma_y)$ and $\mathbb{B} \xrightarrow{d} N(0, \sigma_{y^*})$ for some finite σ_y, σ_{y^*}. Then by the Mann-Wald Continuous Mapping Theorem and the replication property of the normal distribution, $(\mathbb{E}[y_i^*]\mathcal{A} - \mathbb{E}[y_i]\mathcal{B})$ is asymptotically normal. Finally, by the Slutsky Transformation Theorem, the product of this term and $\mathbb{C}$ is also asymptotically normal.

A.4 More Details on Estimating the Production Function

This appendix explains why I use the Anderson-Hsiao estimator instead of fixed-effects and shows measurement error in the factors of production might not bias my estimates of the production function too badly under plausible assumptions.

A.4.1 Fixed-Effects Versus Anderson-Hsiao

Farmers have individual invariant productivity terms, so the standard approach is to estimate the production function

$$y_{it} = A_{it}\phi_{it}K_{it}^{\alpha}T_{it}^{\beta}L_{it}^{\lambda}$$

in logs using fixed-effects (the within estimator). In other words, sub (1.6) and (1.7)

apply the Delta method to prove the asymptotic normality of $\hat{\mathbb{G}}_{\hat{I}_t}$.

into the logged production function, and estimate by OLS

$$\log y_{it} = [Household\ Fixed\ Effect]_i + a^H [Hunger]_{it} + \sum_j a_j^S [Dummy\ Shocks]_{j,it}$$
$$+ \sum_k a_k^D [District - Year\ Dummies]_{k,it} + \sum_m a_m^R [Monthly\ Precipitation]_{m,it}$$
$$+ \theta_K \log K_{it} + \theta_T \log T_{it} + \theta_L \log L_{it} + [Overall\ Error]_{it}. \tag{A.6}$$

However, the consistency of the estimates requires what Wooldridge (2002) calls the strict exogeneity assumption:

$$\mathbb{E}[[Overall\ Error]_{is} \mid K_{it}, T_{it}, L_{it}] = 0 \qquad \forall s, t = 1, \ldots, T$$

The unexplained productivity component $[Overall\ Error]_{it}$ must be completely cross-temporally uncorrelated with the regressors of interest, in particular the productive factors.

To see why this might be problematic, consider a farmer who gets a bad unobserved shock because he accidentally used too much fertilizer. Then although the accident will not affect contemporaneous or past factor input choices (because it is unpredictable), it might affect future input choices if the resulting low yield drives him to sell land or capital for food. A drop in the possession of these factors would be informative about an earlier bad productivity shock. If the unobserved shock is big relative to household income then OLS will estimate production shares inconsistently.

Instead I make a weaker assumption of sequential exogeneity Wooldridge (2002). Sequential exogeneity only requires that factor levels be uncorrelated with future values of unexplained productivity:

$$\mathbb{E}[[Overall\ Error]_{is} \mid K_{it}, T_{it}, L_{it}] = 0 \qquad \forall t = 1, \ldots, T \text{ and } s \leq t$$

A bad shock last year can affect capital stock this year as in the previous example without damaging consistency. Problems only arise if farmers anticipate some part of the error term, like a technological innovation they acquire but I do not observe. If farmers buy more land or capital to exploit the innovation the estimator will not be consistent. Large technological innovations are unlikely in Thailand and affect everyone, so the time and time-district dummies should capture them. Assuming sequential exogeneity is the same as assuming $\hat{A}_{it}$ completely captures the anticipated component of TFP. I must assume as much, anyways, for the optimal allocation I compute to be accurate.

To implement the assumption I first-difference away the fixed component and instrument the differenced factors with their lagged levels. The first stage of the regression for capital is

$$\Delta \log K_{it} = +b^H \Delta[Hunger]_{it} + \sum_j b_j^S \Delta[Dummy\ Shocks]_{j,it} \tag{A.7}$$
$$+ \sum_k b_k^D [District - Year\ Dummies]_{k,it} + \sum_m b_m^R \Delta[Monthly\ Precipitation]_{m,it}$$
$$+ \nu_K \log K_{it} + \nu_T \log T_{it} + \nu_L \log L_{it} + \Delta[Overall\ Error]_{it}.$$

and the second stage is

$$\Delta \log y_{it} = a^H \Delta[Hunger]_{it} + \sum_j a_j^S \Delta[Dummy\ Shocks]_{j,it} \tag{A.8}$$
$$+ \sum_k a_k^D [District - Year\ Dummies]_{k,it} + \sum_m a_m^R \Delta[Monthly\ Precipitation]_{m,it}$$
$$+ \theta_K \Delta \log K_{it} + \theta_T \Delta \log T_{it} + \theta_L \Delta \log L_{it} + \Delta[Overall\ Error]_{it}.$$

The identification assumption is

$$\mathbb{E}[(\log K_{i,t-1})[Overall\ Error]_{i,t-1}] = \mathbb{E}[(\log K_{i,t-1})[Overall\ Error]_{i,t}] = 0$$

and so on for land and labor, as well.

A.4.2 Measurement Error in Factor Choices

Given how I construct the measures of capital and labor, measurement error is likely. But under some assumptions about the nature of the error it should not bias estimates too badly.

Suppose the error ς follows an autoregressive process, relates the true value to the observed value multiplicatively, and is independent of the true value. For simplicity, focus on capital and ignore TFP modifiers as well as labor and land; this simplification should not qualitatively affect the argument. Formally, let K_{it}^o be the observed value of capital and suppose

$$\log y_{it} = c_i + \alpha \log K_{it} + \varepsilon_{it}$$
$$K_{it} = \exp(\varsigma_{it}) K_{it}^o$$
$$\varsigma_{it} = \rho \varsigma_{i,t-1} + \xi_{it}$$

Assume that ρ is close to 1 and ξ_{it} is small; that is, the error is highly persistent over time. This assumption is plausible because of the way I calculate factor stocks. A household's level of capital, for example, is computed by backward depreciating the most recent purchase of each asset type and aggregating. Any error introduced by associating latest purchase price with productive value would be carried through to previous years.

The primary regression equation is then

$$\begin{aligned}
\Delta \log y_{it} &= \alpha \Delta \log K_{it} + \Delta \varepsilon_{it} \\
&= \alpha \Delta \log K_{it}^o + \Delta \varepsilon_{it} - \alpha(1-\rho)\varsigma_{i,t-1} + \xi_{it}.
\end{aligned}$$

The first-stage regression is

$$\begin{aligned}
\Delta \log K_{it} &= \pi_K \log K_{i,t-1} + \epsilon_{it} \\
\Rightarrow \Delta \log K_{it}^o &= \pi_K \log K_{i,t-1}^o + \epsilon_{it} + \pi_k \varsigma_{i,t-1}.
\end{aligned}$$

The estimate of π_k will be attenuated because of the correlation between $K_{i,t-1}^o$ and the lagged measurement error terms. But this will not matter as long as the first-stage remains strong, and the results suggest it is.

More important is that the lagged level of capital is correlated with $-\alpha(1-\rho)\varsigma_{i,t-1}+\xi_{it}$. But under the assumptions, this lagged error term is small (because $1-\rho$ and ξ are small), so the bias of the IV estimator will be small as long as the first-stage is strong.

The intuition is simple: when the measurement error is persistent, it is captured by the fixed-effect so the parameter estimates are unbiased. What then happens to the estimates of efficiency when these fixed-effects make up some portion of anticipated TFP? Assuming the measurement error is independent of the true fixed component of anticipated TFP, it will almost certainly bias estimated misallocation upwards. To see why, assume the extreme case where markets are complete and contingent, so the observed outcome is perfectly efficient. My noisy estimates of TFP will produce a noisy computation of the efficient outcome; this will differ from the realized outcome, so the measure will erroneously mark the perfectly efficient allocation as inefficient. The true misallocation in rural Thailand may be even lower than I find.

Although measurement error in capital probably will be very persistent, that need not hold for labor. If a household member changes his actual time in the fields often the error in measured labor might fluctuate from year to year. The reader should keep that caveat in mind while evaluating the results.

A.5 Specification Tests

This appendix runs simple tests of whether the production function is approximately isoelastic (Cobb-Douglas) and whether my measure of anticipated productivity actually captures the farmer's ability to farm.

A.5.1 Is the Production Function Isoelastic?

The Cobb-Douglas production function is a special case of the class of constant-elasticity production functions where the elasticity of substitution ϵ between factors is 1 (hence alternative label "isoelastic"). I follow the procedure in Udry (1996) and suppose

$$y_{it} = A_{it}\phi_{it}\left[\alpha K_{it}^{\frac{\epsilon-1}{\epsilon}} + \beta T_{it}^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha-\beta)L_{it}^{\frac{\epsilon-1}{\epsilon}}\right]^{\sigma\frac{\epsilon}{\epsilon-1}} \tag{A.9}$$

where σ denotes the returns to scale. For computational simplicity I assume $A_{it} = A_i e^{\vartheta t}$, the product of a fixed effect and a time trend. Take logs of both sides and subtract away the within household mean to eliminate the fixed-effect. [3]

Column 1 of Table A.1 reports the results of estimating the transformed equation with nonlinear least squares. The test of interest is whether ϵ differs substantially from 1. As (1) of Table A.1 indicates, this null is actually rejected. However, the point estimate is almost identical to one ($\hat{\epsilon} = 1.013$) and rejection occurs mainly because the variance of these estimates is very small. The envelope theorem guarantees misallocation does not change much with small changes in the elasticity of substitution, so a tiny deviation from Cobb-Douglas production should not change the results much.

[3]The ideal way to estimate the equation is the nonlinear equivalent of Anderson-Hsiao: applying GMM to the first-differenced form of (A.9) using legged factors as instruments. Unfortunately, it does not converge.

Table A.1
CES Production Function Estimates

	(1)
	NLS
	b/se
ϵ	1.013
	(0.0068)
σ	0.767
	(0.0408)
α	0.311
	(0.0086)
β	0.369
	(0.0119)
ϑ	0.000
	(0.0000)
N	775.000
NT	6230.000
Pval: $\epsilon = 1$	0.061

Note: Estimated using fixed-effects nonlinear least-squares. Standard errors are bootstrapped with resampling at the household-level.

A.5.2 Is A_{it} Really a Valid Measure of Productivity?

Anticipated productivity A_{it} almost wholly determines a farmer's allocation with perfect markets. If $\hat{A}_{it}$ is a bad measure of true productivity, the optimal allocations I compute could be completely wrong. If $\hat{A}_{it}$ actually measures a farmer's productivity it should be correlated with individual and village characteristics that make a farmer more productive.

For example, if $\hat{A}_{it}$ is capturing actual productivity it should be higher in areas with climate and soil better suited to growing rice. And if years spent in school raise productivity, it should also be higher for more educated farmers. Table A.2 shows a regression of anticipated TFP $\hat{A}_{it}$ on province dummies and two proxies: an index of agro-climactic suitability for rice and the farmer's schooling. For schooling I focus on primary education, as basic skills like literacy matter most in farming. The results indicate that TFP is indeed strongly positively correlated with climactic suitability, and also correlated with the farmer's education.

Table A.2
Partial Correlations of Anticipated TFP

	(1)
	log $\hat{A}$
	b/se
Primary Schooling (years)	0.058***
	(0.02)
Beyond Primary	0.026
	(0.08)
Rice Suitability	0.082***
	(0.02)
N	775.000
NT	6216.000

Note: The regression shows the correlation between predictors of productivity and my estimates of anticipated productivity. I cluster standard errors by household and include province dummies. Primary Schooling (years) represents the number of grades of primary school the household head has completed, where someone who has completed any grades in secondary school (Matthayom) is assumed to have completed six years of primary school (Prathom). Beyond Primary is a dummy for additional schooling, and Rice Suitability a district-level agro-climactic index.

A.6 Data Appendix

A.6.1 Village (or higher) Level Variables

International Rice Prices From the IMF's commodity price data. I took the yearly average.

Village Wage Rates From Section V of the annual household survey.

For 1996: For each household, find any worker in the "other" category who lists their occupation as related to "labor" or "labour" and compute their daily wage. Construct medians by village, subdistrict, etc.

For 1997-2008: For each household, find any worker listed as general agricultural laborer of any sort or in the "other" category reporting an occupation related to "labor" or "labour" and compute their daily wage. Construct medians by village, subdistrict, etc.

Village Population From Section iii of the annual key informant survey. Survey records both number of households and population of the village.

Precipitation I obtained gridded monthly rainfall estimates to cover Thailand from 1996 through 2008. The estimates for 1996 and 1997 were .5 x .5 lat-long degree grids from the University of Delaware Climate Project's Terrestrial Precipitation Gridded Monthly Time Series. Those for the rest of the year were .25 x .25 lat-long degree grids from NASA's Tropical Rainfall Measuring Mission (Product 3B-43). I used ArcGIS's Topo-to-Raster tool to create an interpolated raster for the rainfall data. I then used district-level boundaries from the Global Administrative Areas Project (GADM) to construct district-level monthly averages. I converted the levels to fractional deviations from the mean for each month (where the monthly mean was computed over the sample period). The rainfall shocks relevant for a particular year match the survey response period (so rainfall in 1996 is rainfall from May 1996 through April 1997).

Rice Suitability Index I obtained the Food and Agriculture Organization's Global Agro-Ecological Zone (GAEZ) data for the climactic suitability of rice and maize. The data for rice suitability came from Plate 38: Suitability for rain-fed and irrigated Rice (high input). The data for maize suitability came from Plate 30: Suitability for rain-fed Grain Maize (intermediate inputs). I computed a zonal mean for each district: an average value indexing the climate suitability of the district for each crop. I inverted the index so higher values correspond to greater suitability.

A.6.2 Household-Level Variables

A.6.2.1 Factors of Production

Land I use the land cultivation data from Section XIV (Landholdings) of the Annual Household Survey. Households report the quantity and value of land they cultivate (regardless of ownership) by use; that is, they separately report land for rice, field crops, orchards, and vegetables. I total the area of the plots for each use and mark this as the land cultivated for each crop. I also deflate the reported value of the plots and total for each crop to form the value of the land owned.

Capital *Owned Mechanical Capital:* I use Section XII (Agricultural Assets) of the Annual Household Survey. The survey reports the number of assets of each type, where I group the following assets into broad categories by depreciation rates: tractors (walking tractors, large and small four-wheel tractors), machines (sets, sprinklers, and threshing machines) and structures (crop storage buildings). I depreciate tractors like vehicles, so the depreciation rates I use are 2 percent for structures, 10 percent for machines, and 20 percent for tractors. I correct clear errors in the series of asset classes where an asset disappears and reappears without any record of a sale or appears and disappears without any record of a purchase. I then construct the value of assets owned at the beginning of the first survey round by deflating and depreciating the purchase price by the year of purchase. I then attempt to follow each asset over time, where I label a piece of equipment a separate asset if the quantity of a certain type of equipment rises from zero to some positive number. I unfortunately must treat the addition of new pieces of equipment to an existing stock as identical to the existing assets of that class; but it is fairly rare that a household has more than one piece of equipment of a certain class. I then assign a ”price” to each asset with the sale value at the very latest transaction date I can find for it (where the initial value in the first survey round is also considered a transaction). I adjust that price for deprecation in preceding and following years and compute the asset value in a given year by multiplying the price by the quantity held. [Recall the quantity is almost always one if the household owns any.] If I cannot identify a price, I am forced to drop the asset from my calculations. [In rare cases where I can identify a year of acquisition but not a price, I use the intertemporal median of village, subdistrict, or district medians for the equipment type.] I then aggregate the value of all assets for each household in each year to construct the value of owned mechanical capital.

Buffalos: I assume buffalo are the only animal used to harvest rice and compute the value of buffalo using the appropriate responses from Section XII (Agricultural Assets) of the Annual Household Survey. The household reports the total current value of all buffalos owned, which I deflate. Missing values for this variable generally mean the question does not apply (e.g. the household owns no buffalo), so I treat missing values as zero.

Capital Expenses: For rented capital, maintenance expenses, and intermediate inputs (which I treat as capital) I use the portion on farm expenses in Section XVI (Income) of the Annual Household Survey. After deflating all currency, I compute intermediate inputs as the sum of expenses on seeds, fertilizers, pesticides/herbicides, and fuel. I then rescale the value of rented capital by a user cost: the depreciation rate plus an interest rate, which I set as 4 percent in line with the literature. It may seem strange to assume a common interest rate given the possibility of financial market failures; but recall my objective is to create a consistent measure of the productive value of the capital owned. Allowing the productive value of a tractor to vary based on the household’s borrowing cost makes no sense. I do not know how much of the rental cost goes to machines versus vehicles, so I take the depreciation rate as the average of the rates for each type of asset. Finally, I add the value of maintenance expenses, which is investment (recall I assume investment is immediately productive).

Total capital is the sum of the value of owned mechanical capital, buffalos, and capital expenses.

Labor *Family Labor:* I first construct the quantity of family labor using Section V (Occupation) of the Annual Household Survey. In each year I count the number of household members who report being unpaid family laborers in their primary occupation and report farming of any sort as their primary or secondary occupation (or report working in the "FIELDS" if their occupation is not categorized). (Some farmers grow several types of crops, but the survey only allows two responses for occupation. To deal with this problem, I reason that a household growing rice will use its working family members on all of its fields, so any family member who works in the fields necessarily works in rice.) I define the number of family workers as the number of household members who satisfy this criterion. I have no intensive margin information on how much the household works, so I assume all members work sixty days of the year in the fields (the median number of days worked from the two years of the Monthly Household Survey available at the time of writing). I aggregate the per-member days worked for each household-year to compute the quantity of family labor. (In other words, I multiply the number of family workers by the median days an individual works on their fields conditional on working at all.)

Hired Labor: The only measure of hired labor is the expenditure on wages recorded among the farm expenses in Section XVI (Income) of the Annual Household Survey. I divide the total expenses on wages by the village-level median daily wage (see above) to construct a measure of days worked by hired labor. Total labor is simply the sum of family and hired labor.

A.6.2.2 Productivity Modifiers

Catastrophes/Bad Income Shocks I use the questions about bad income years from Section II (Risk Response) of the Annual Household Survey. The household reports the worst of the last several years for income (including the response year), and the reason for it being atypically bad. If a household chooses the response year as the worst, I mark it as suffering one of the following catastrophes based on the reason it gives:

- Reports bad income this year due to illness
- Reports bad income this year due to death in family
- Reports bad income this year due to retirement
- Reports bad income this year due to flooding

- Reports bad income this year due to crop-eating pests
- Reports bad income this year due to poor rainfall
- Reports bad income this year due to low yield for other reasons
- Reports bad income this year due to low price for output

Hunger/Undernourishment I have no direct measure of calories and instead adapt the notion of the staple budget share (SBS) introduced by Jensen and Miller (2010). I use consumption expenditure data from Section XV (Expenditure) to compute the fraction of the household's budget spent on the staple food in Thailand: rice. This measure includes the value of rice the household grew itself. The intuition behind this measure is that as a household becomes wealthier (and less hungry) it substitutes away from the staple crop towards other foods (which are superior goods). The higher the SBS, the more likely it is the household is hungry.

A.6.2.3 Other Variables

Revenue I use the questions about gross income from Section XVI (Income) of the Annual Household Survey. Households report their revenue from each of several sources, including rice farming and other agricultural activities. Enumerators explicitly reminded households to include the value of crops they produced and then consumed. I deflated and constructed income variables for each of the following sources: Rice Farming, Corn Farming, Vegetable Farming, Orchard Farming, Other Farming.

Education I use the questions about age and education from Section IV (Household Composition) of the Annual Household Survey. I keep information about the age, highest grade completed, and school system of the household head. I defined separate variables for number of years spent in primary school (generally from 1-6 for P1-P6), number of years spent in secondary school (generally 1-6 for M1-M6, unless the individual chose the vocational rather than academic track, in which case I set years of secondary school to 3), years of vocational school (from 1-3 for PWC1-PWC3, PWS1-PWS3, or PWT1-PWT3), and years of university (from 1-4).

Rice Farming Experience Households report the number of years spent at their primary occupation in Section V (Occupation) of the Annual Household Survey. I record the years spent for individuals who report rice farming as their primary occupation and categorize themselves as ”owners” of the business. I take this as a measure of the rice-farming experience of the household (head). In the rare cases where multiple household members claim to be the owner of a rice-farming business, I take the median as the household-level experience.

Constraints I use the questions on farm expansion from Section XII (Agricultural Assets) of the Annual Household Survey. A household is labeled as ”constrained” if it reports there is room for profitable expansion in its business. I label it credit-constrained if it reports insufficient money for labor, land, or equipment among the reasons for not expanding. I label it factor-constrained if it reports not enough land or labor (a distinct response from insufficient money) among the reasons for not expanding. A household can be both credit- and factor-constrained. I further label households as exclusively credit- or factor-constrained if they report a constraint in one but not the other.

Risk-Aversion In 2003 the survey started posing to households a hypothetical choice between staying at their current income forever and taking a job that with 50-50 chance pays either double or two-thirds their current income. If they choose their current job the interviewer gives them the same choice except the alternate job now has a 50-50 chance of paying either double or 80 percent of their current income. If the respondent chose his current job for both questions I marked it as having “high risk aversion.” Since the question was not asked in 1996, I use the 2003 question in Section 1.6.

Savings I use Section XIX (Savings) of the annual household survey. I take the total savings each household has deposited with commercial banks, agricultural cooperatives, the Bank for Agriculture and Agricultural Cooperatives, PCG village funds, and a rice bank.

APPENDIX B

Appendix: Risky Income or Lumpy Investments?

B.1 Proofs

B.1.1 Generalizing the Risk and Return Predictions

Letting $M \in \{1, 2, \ldots\}$, the optimal labor allocation is

$$L_p = \frac{\bar{w}_+ M + \alpha\sigma_s^2}{\alpha\left(M\sigma_p^2 + \sigma_s^2\right)} \tag{B.1}$$

Consider the threshold fixed cost that separates households who choose M activities from those who choose $M+1$ activities:

$$-e^{-\alpha\bar{C}(M)+\frac{\alpha^2}{2}V(M)} = -e^{-\alpha\bar{C}(M+1)+\frac{\alpha^2}{2}V(M+1)}$$

The threshold is

$$\bar{F}_M = \frac{\sigma_s^2\left(\alpha\sigma_p^2 - \bar{w}_+\right)^2}{2\alpha\left(M\sigma_p^2 + \sigma_s^2\right)\left((M+1)\sigma_p^2 + \sigma_s^2\right)}.$$

The derivatives with respect to σ_p^2 and $\bar{w}_p$ are

$$\frac{\partial\bar{F}_M}{\partial\sigma_p^2} = \frac{\sigma_p\sigma_s^2\left(\alpha\sigma_p^2 - \bar{w}_+\right)\left(\alpha\sigma_s^2\left((2M+1)\sigma_p^2 + 2\sigma_s^2\right) + \bar{w}_+\left(2M\left((M+1)\sigma_p^2 + \sigma_s^2\right) + \sigma_s^2\right)\right)}{\alpha\left(M\sigma_p^2 + \sigma_s^2\right)^2\left((M+1)\sigma_p^2 + \sigma_s^2\right)^2}$$

$$\frac{\partial\bar{F}_M}{\partial\bar{w}_p} = -\frac{\sigma_s^2\left(\alpha\sigma_p^2 - \bar{w}_+\right)}{\alpha\left(M\sigma_p^2 + \sigma_s^2\right)\left((M+1)\sigma_p^2 + \sigma_s^2\right)}$$

Since by assumption $\left(\alpha\sigma_p^2 - \bar{w}_+\right) > 0$, $\frac{\partial \bar{F}_M}{\partial \sigma_p^2} > 0$ and $\frac{\partial \bar{F}_M}{\partial \bar{w}_p} < 0$ for all M. Then a rise in the riskiness of the primary activity will cause all the thresholds to rise, meaning households will be willing to pay more for any number of activities. This will cause the average number of activities in the sample to rise. A parralel argument shows a rise in the average return decreases all thresholds and decreases the average number of activities.

QED

B.1.2 Verifying the Cost Prediction

We can rewrite expected side income as

$$\begin{aligned}
\mathbb{E}[y_s] &\approx \bar{w}_s - F\left(\frac{\partial M}{\partial \sigma_p^2}\sigma_p^2 + \frac{\partial M}{\partial \bar{w}_p}\bar{w}_p\right) + \bar{w}_s\left(-\frac{\partial L_p}{\partial M}\left[\frac{\partial M}{\partial \sigma_p^2}\sigma_p^2 + \frac{\partial M}{\partial \bar{w}_p}\bar{w}_p\right] - \frac{\partial L_p}{\partial \sigma_p^2}\sigma_p^2 - \frac{\partial L_p}{\partial \bar{w}_p}\bar{w}_p\right)\\
&= \bar{w}_s + \left(-F - \bar{w}_s\frac{\partial L_p}{\partial M}\right)\left[\frac{\partial M}{\partial \sigma_p^2}\sigma_p^2 + \frac{\partial M}{\partial \bar{w}_p}\bar{w}_p\right] + \bar{w}_s\left[-\frac{\partial L_p}{\partial \sigma_p^2}\sigma_p^2 - \frac{\partial L_p}{\partial \bar{w}_p}\bar{w}_p\right]\\
&= \bar{w}_s + \left(-F - \bar{w}_s\frac{\partial L_p}{\partial M}\right)\hat{M} + \bar{w}_s\left[-\frac{\partial L_p}{\partial \sigma_p^2}\sigma_p^2 - \frac{\partial L_p}{\partial \bar{w}_p}\bar{w}_p\right]\\
\rightarrow y_s &= \bar{w}_s + \left(-F - \bar{w}_s\frac{\partial L_p}{\partial M}\right)\hat{M} + \eta + \varepsilon
\end{aligned}$$

where $\hat{M}$ is the predicted number of activities from the first-stage regression, η is the direct effect of labor reallocation from changes in the volatility and average returns to the primary activity, and ε is an independent error term.

The instrumental variables estimate is consistent for the value

$$\gamma_A = -F - \bar{w}_s\frac{\partial L_p}{\partial M} + \mathbb{E}[\hat{M}\gamma]$$

If $-\bar{w}_s\frac{\partial L_p}{\partial M} + \mathbb{E}[\hat{M}\gamma] > 0$, then $\gamma_A > -F$, which implies that if $\gamma_A < 0$ then $-F < 0$ and thus under-specialization is costly.

First I show that $\frac{\partial L_p}{\partial M} < 0$. From the expression for L_p found in (B.1) in Appendix B.1.1 we have that a 1 unit increase in M will cause a rise in the numerator of $\bar{w}_+$ and a rise in the denominator of $\alpha\sigma_p^2$. Since by assumption $\alpha\sigma_p^2 > \bar{w}_+$, the denominator rises by more than the numerator and the total effect is negative. Thus, $\frac{\partial L_p}{\partial M} < 0$.

Now I show that $\mathbb{E}[\hat{M}\gamma] > 0$. The expectation equals

$$
\begin{aligned}
\mathbb{E}[\hat{M}\gamma] &= +\bar{w}_s \left[-\frac{\partial L_p}{\partial \sigma_p^2}\mathbb{E}[\hat{M}\sigma_p^2] - \frac{\partial L_p}{\partial \bar{w}_p}\mathbb{E}[\hat{M}\bar{w}_p] \right] \\
&= -\bar{w}_s [\underbrace{\frac{\partial L_p}{\partial \sigma_p^2}}_{-} \underbrace{\frac{\partial M}{\sigma_p^2}}_{+} + \underbrace{\frac{\partial L_p}{\partial \bar{w}_p}}_{+} \underbrace{\frac{\partial M}{\bar{w}_p}}_{-}]
\end{aligned}
$$

where the final equality applies the definition of $\hat{M}$; applies the predictions of the effects of risk and returns to the number of activities to get the signs of $\frac{\partial M}{\sigma_p^2}$ and $\frac{\partial M}{\bar{w}_p}$; and takes the deriviatives of L_p found in (B.1) in Appendix B.1.1 with respect to σ_p^2 and $\bar{w}_p$. This proves that $\mathbb{E}[\hat{M}\gamma] > 0$.

QED

B.2 Detailed Data Appendix

B.2.1 Time Series Variables

- **Consumer Prices:** From Bank of Thailand monthly index, acquired from Global Financial Data database. Data were used with permission of Global Financial Data.

- **International Rice Price:** Acquired from IMF monthly commodity price data. Deflated using monthly consumer price index.

B.2.2 Panel Variables

- **Rice Harvest:** From module 7 (Crop Harvest) section of the monthly survey. Keep only un-milled rice (both sticky and non-sticky). Define rice harvest soon as a reported positive harvest of unmilled rice in the subsequent three months. Define rice harvest past as having had positive harvest of unmilled rice in the current or previous three months. Define rice farmer (or rice harvest ever) as having had a positive rice harvest at any point in the survey span.

- **Crop-Plots:** From module 5 (Crop Activities) section of the monthly survey. Make the monthly aggregate of "value transacted" for each households sale of each crop. This is the revenue from crops. For number of crop plots, I use the "projected harvest" table, which asks farmers to predict revenue for each productive crop. Every entry corresponds to a different perceived revenue stream for the farmer, so I take number of crop-plots as simply the count of these for each household in each month.

- **Aquaculture:** From module 10 (Fish-Shrimp) of the monthly survey. For each household, make monthly aggregates of the value of fish and shrimp output; this is the revenue from aquaculture. I compute whether a household does aquaculture as whether it reports raising fish/shrimp or having shrimp ponds in a given month.

- **Large Businesses:** From module 12 (Household Business) of the monthly survey. For each household, make monthly aggregates of the cash and in-kind revenue plus the value of products/services consumed by the household; this is the revenue from large businesses. Compute the number of businesses for each household as the number of entries in the household report of revenues.

- **Small/Miscellaneous Businesses:** From module 24 (Income) of the monthly survey. For each household, make monthly aggregates of the cash and in-kind revenue for each "other" income source; this is the revenue from miscellaneous businesses. Compute the number of miscellaneous activities for each household as the number of entries in the household report of revenues.

- **Number of Jobs:** From module 11 (Activities-Occupation). For each person and each job number in any month, mark if it was worked the previous two and the following two months (note that jobs are not assigned job numbers in their first months, so technically I only check the previous one month as it must have been worked the month before to have an ID). If so, it is a "steady job." I count each households total number of jobs and steady jobs each month, then compute the number of unsteady jobs as the difference. For each job and each month, sum the cash and in-kind payments and aggregate by household-month. This is the monthly job revenue.

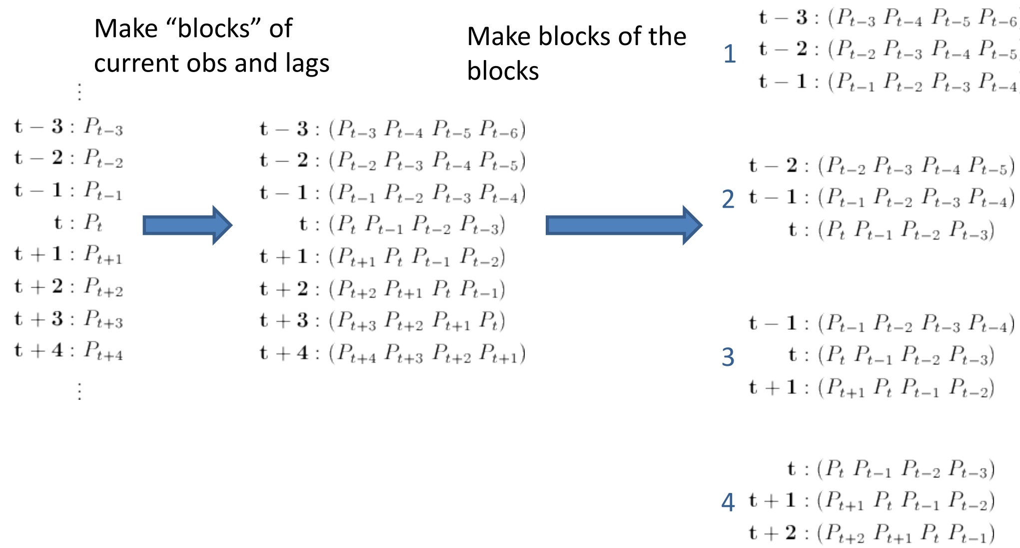
- **Number of Activities:** I define number of activities as simply the sum of the number of crop plots, the number of livestock activities, the indicator for practice of aquaculture, the number of large businesses, the number of jobs, and the number of miscellaneous activities.

- **Total Revenue, Consumption, and Transfers:** Total revenue is the sum of revenue from crop activities, livestock activities, aquaculture, large businesses, jobs, and miscellaneous activities. Total consumption is the sum of all domestic expenditures by both cash and credit plus consumption of home-produced goods. Expenditures reported at a weekly rather than monthly frequency (in module 23W, Weekly Expenditures Update) are aggregated by month for each household and added to those reported at a monthly frequency (in module 23M, Monthly Expenditures Update). Transfers are defined as the household's net incoming transfers. More precisely, I aggregate by household-month the transfers from people inside and outside the village and subtract similarly aggregated transfers to people inside and outside the village (all found in module 13 on Remittances). I use only transfers not earmarked for a specific event because these unplanned transfers are more like insurance.

B.3 Inference: The Two-Stage Bootstrap

The predicted mean and volatility are both generated regressors, so I must adjust my inference to account for their presence. It is easy to see that under my assumptions the full estimators match the conditions for Murphy and Topel (2002). Directly applying their analytic expressions is inconvenient and also problematic because small sample bias in the

Figure B.1
Bootstrap, Step 1: Forming Blocks of Blocks



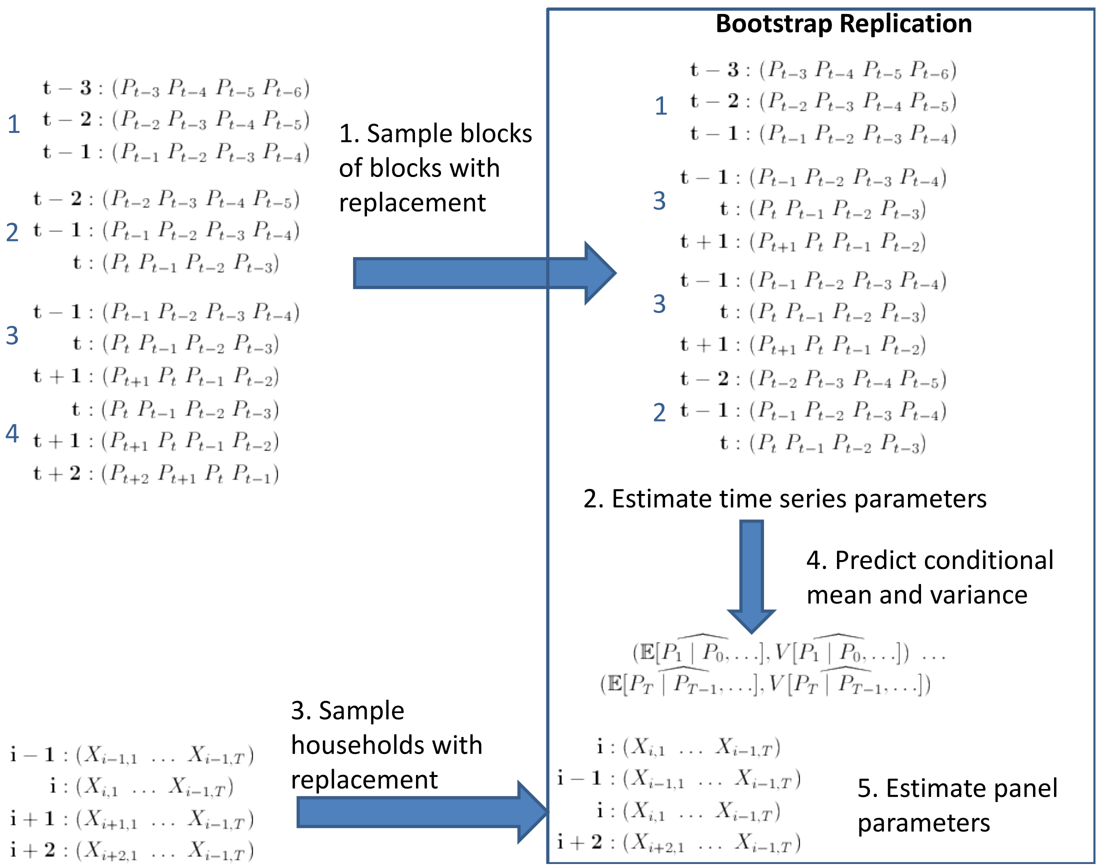
Note: First, I prepare the time series of rice prices for resampling. I form "blocks" consisting of the current price and however many lags I need to estimate the time series model. I then group every observation into one or more "blocks of blocks," adjacent interlocking sets of observations and their associated lags.

time series estimates might produce an abnormal small sample distribution for the estimated parameters. But the asymptotic normality their propositions guarantee also ensures the validity of bootstrapped confidence intervals and hypothesis tests.

I implement the procedure as outlined in Figures B.1-B.3. First, I prepare the time series of rice prices for resampling. I form "blocks" consisting of the contemporaneous price and however many lags I need to estimate the time series model. I then group every observation into one or more "blocks of blocks," contiguous interlocking sets of observations and their associated lags.

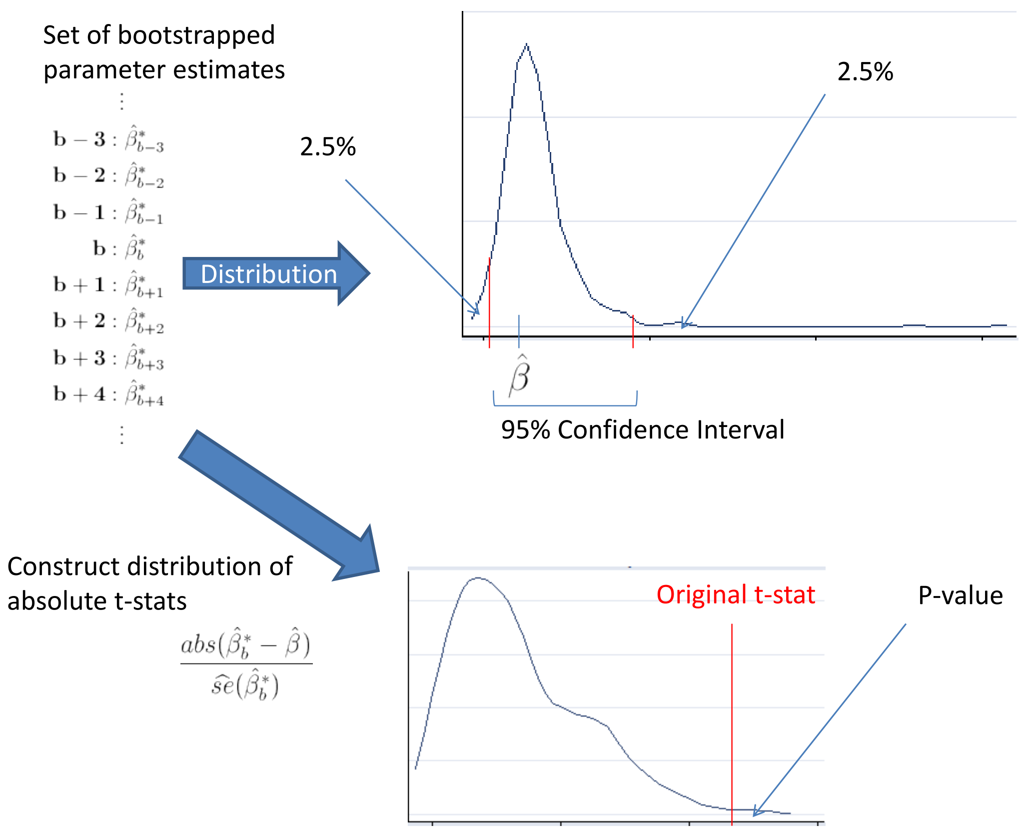
Next, I run the bootstrap replications. Each replication follows five intermediate steps. First, I sample with replacement the blocks of blocks of rice prices to construct a bootstrapped time series of equal length to the original time series. I estimate the parameters of the time series model on the bootstrapped data. I then resample with replacement house-

Figure B.2
Bootstrap, Step 2: Bootstrap Replications



Note: Next I run the bootstrap replications. Each replication follows five steps. First I sample with replacement the blocks of blocks of rice prices to construct a bootstrapped time series of equal length to the original time series. Next I estimate the parameters of the time series model on the bootstrapped data. I then resample with replacement households (together with all their monthly observations) from the panel to construct a bootstrapped panel with as many households as the original panel. Then I use the estimated time series model to predict the conditional mean and variance of the international rice price for each household-month observation. Finally, I estimate the panel specification and record the resulting coefficients. I run 2000 replications for the risk specifications and 3000 replications for the IV specifications.

Figure B.3
Bootstrap, Step 3: Constructing Confidence Intervals and P-Values



Note: I construct confidence intervals with the dataset of bootstrapped parameters to find the 2.5th and 97.5th percentiles. These are the boundaries of the 95% confidence interval. I then compute the absolute t-statistic centered around the original parameter estimate for each replication. The fraction of these absolute t-statistics that is greater than the original t-statistic is the p-value.

holds (together with all their monthly observations) from the panel to construct a bootstrapped panel with as many households as the original panel. Then I use the estimated time series model to predict the conditional mean and variance of the international rice price for each household-month observation. Finally, I estimate the panel specification and record the resulting coefficients. I run 1000 replications for the risk specification and 2000 replications for the IV specifications.

The final step is to compute confidence intervals and p-values. To construct confidence intervals, I use the dataset of estimated parameters from bootstrap replications to find the 2.5th and 97.5th percentiles. These are the boundaries of the 95% confidence interval. To construct p-values, I compute the absolute t-statistic centered around the original parameter estimate for each replication. The fraction of these absolute t-statistics that is greater than the original t-statistic is the p-value.

B.4 Other Tests of Robustness

B.5 Alternative Model: Minimum Labor Inputs

Is it plausible that the kinds of activities a rice farmer can enter three months before his harvest would, as my model assumes, have a lumpy fixed cost? Finding casual labor or growing cassava may be easy if the farmer has already done so every time prices turned volatile in the past. In this appendix I build a model without fixed costs where risk still causes under-specialization. The prediction's robustness is why I emphasize that my model of risk and under-specialization is not *the* model, but just a convenient tool to formalize the intuition.

Let the household's utility function be as before and for simplicity consider the case of choosing between perfect specialization and one side activity. The household can costlessly enter a side activity but must allocate it at least $\underline{L} > 0$ units of labor. The lower-bound on labor choice captures the idea that it is not worth an employer's time to hire a worker for only a few hours per week, so even work that does not require paying a fixed cost does require a lumpy investment of time. I need the lumpiness to make specialization optimal for some degree of riskiness. Otherwise the household always has a side activity and only varies how much it works on the side activity instead of whether it has one at all. I also assume the average return to the side activity is strictly less than the average return to the primary activity—that is, $\bar{w}^p - \bar{w}^s = w^+ > 0$. The household faces the trade-off

Table B.1
Robustness: Main Results Excluding Pre-Harvest Rice Sales

	(1) Activities	(2) Revenue
	(1)	(2)
	Activities	Revenue
Activities		-14195.18**
		[0.027]
Rice Farmer		
- × Mean	0.00	-86.06
	[0.395]	[0.307]
- × Volatility	-0.09**	-368.73
	[0.011]	[0.531]
Expecting Harvest		
- Main	1.37***	-238.27
	[0.000]	[0.953]
- × Mean	-0.01***	(Excluded Instrument)
	[0.000]	
- × Volatility	0.05*	(Excluded Instrument)
	[0.091]	
Recent Harvest		
- Main	-0.63	-27329.39*
	[0.263]	[0.077]
- × Mean	-0.01***	147.78
	[0.003]	[0.390]
- × Volatility	0.17***	1095.82
	[0.003]	[0.370]
Household Fixed-Effects	Yes	Yes
Month Fixed-Effects	Yes	Yes
Households	743	743
Observations	47395	47395
F-stat Exc. Inst.		10.054
Hansen's J Stat		0.015

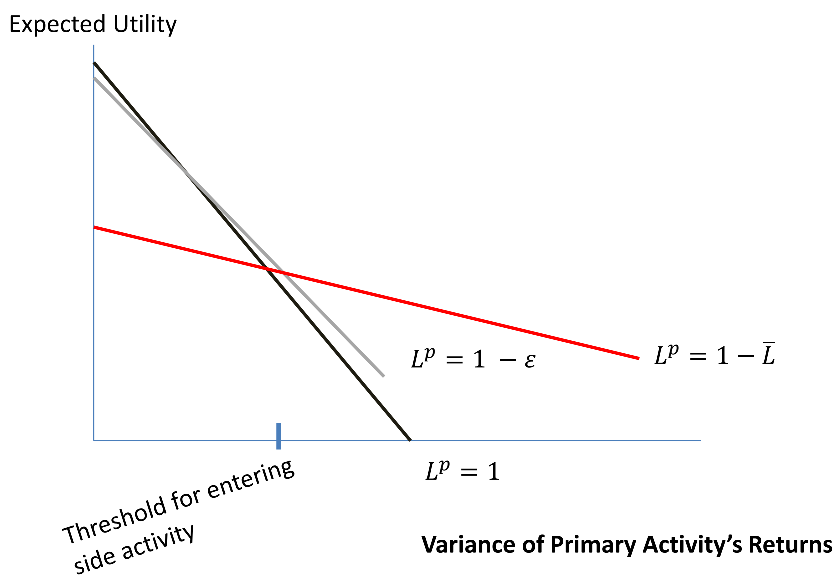
Note: . I exclude observations when households claim they sold rice while still expected their harvest. Volatility still causes households to enter more activities (Column 1) and the extra activities are costly (Column 2).

Table B.2
Robustness: Regressor of Interest Does not Affect Wages

	(1)
	Activities
Mean	0.01
	[0.424]
Volatility	0.66
	[0.167]
Rice Farmer	
- Main	-43.31
	[0.281]
- $\times$ Mean	0.01
	[0.891]
- $\times$ Volatility	0.52
	[0.775]
Expecting Harvest	
- Main	24.62
	[0.478]
- $\times$ Mean	-0.06
	[0.444]
- $\times$ Volatility	-2.90
	[0.483]
Recent Harvest	
- Main	20.65
	[0.193]
- $\times$ Mean	-0.07
	[0.498]
- $\times$ Volatility	-1.80
	[0.477]
Village Fixed-Effects	Yes
Month Fixed-Effects	No
Villages	16
Observations	1152

Note: Suppose wages are correlated with volatility in the rice price. Then the extra jobs the household takes up may not be a response to risk but rather a response to better earnings in side activities. This table demonstrates that median village wages are not correlated with the village averages of any of the regressors of interest. I run the analysis at the village rather than individual level because households might be willing to take lower paying jobs to hedge against risk, lowering the average wage of jobs held even though volatility has no confounding effect on wages.

Figure B.4
Intuition of the Alternative Model



	$M = 0$	$M = 1$
$\bar{C}$	$\bar{w}^p$	$\bar{w}^p - w^+(1 - L^p)$
V	σ_p^2	$(L^p)^2\sigma_p^2 + (1 - L^p)^2\sigma_s^2$

The opportunity cost of the side activity is $w^+(1 - L^p)$, and since it is no less than $w^+\underline{L} > 0$ the household still loses a discrete chunk of expected revenue when it diversifies. Although it does not literally pay a fixed cost the household's trade-off between the mean and variance of consumption is similar to the one it faced in the original model. They are not identical—for example, the cost of diversification is now uncertain—but similar enough for risk to cause under-specialization.

Figure B.4 gives the intuition. With perfect specialization the household's expected utility is maximized when the primary activity's returns have zero variance, but expected utility falls steeply as the variance rises. The household can flatten the utility-variance relationship by moving some labor from the primary activity to the side activity. Without a lower bound on labor devoted to the side activity, the household would always move ε units of labor to the side activity and be happier without perfect specialization. But with a lower bound the household must accept a discretely lower and flatter utility-variance relation. If the variance of the side activity is low, the household prefers specialization. But when the variance exceeds a critical threshold the household prefers to diversify. If w^+ has a nondegenerate distribution the average number of activities will rise continuously with the variance. Then the lower bound model makes the same prediction $\frac{d\mathbb{E}[M]}{d\sigma_p^2} > 0$ as the fixed cost model from the main text.

APPENDIX C

Appendix: Does Factionalism Distort Production?

C.1 Additional Tables

Table C.1
Bribes by Official

Official	Prob. Bribe	Avg. Bribe	% All Bribes	% Amount
Lekhpal/Patwari	59.89	118.38	18.56	15.09
Line-Man - Electricity Department	43.75	89.12	15.74	9.54
Village Secretary	20.89	105.24	8.82	6.31
PDS - Functionary	9.58	90.50	5.87	3.63
Village Development Officer	51.54	76.46	4.76	2.47
Auxiliary Nurse Midwife (ANM)	17.89	117.84	4.24	3.40
Doctor - Local Animal Husbandry Centre/Hospital	27.93	124.38	3.80	3.24
Gram Pradhan	5.27	346.62	3.59	8.54
Pharmacist at PHC/Community Health Centre	10.01	93.86	3.19	2.04
Medical Officer (MOIC)/Govt. Doctor	15.11	129.33	2.84	2.51

Table C.2

Bribes by Service

Service	Prob. Bribe	Avg. Bribe	% All Bribes	% Amount
To get a copy of the RoR	55.03	140.19	10.98	10.56
Ration Cards (APL/BPL/Antodaya)	20.89	71.11	9.74	4.73
For maintenance & repair of street lights, wires, transformers, etc.	43.24	126.64	6.88	5.91
To get a new electricity connection	61.07	86.91	6.14	3.62
Proof of Address Verification	37.54	73.45	6.09	3.05
To get free medicines & immunization of cattle under specific Government schemes	28.58	102.50	3.83	2.70
For obtaining caste certificate	65.47	142.79	3.46	3.37
To get a copy of the household details from the Kutumb/Family Register	50.32	67.30	3.37	1.56
For getting additional benefits under the PDS system	21.28	97.17	3.08	2.05
For rectification of electricity bills	26.81	130.22	2.85	2.53

Table C.3

Evidence of Distortions to Production

			A. Services				
	Gave Bribe	Total Services	Total Borrowed	Average Interest	Loan from Govt.		
Estimate	-0.937***	1.992	-44.966	4.803	0.026		
	(0.336)	(3.017)	(703.018)	(10.291)	(0.200)		
Observations	1054	1054	1054	486	1054		
Elections	88	88	88	70	88		
			B. Production				
	Gave Bribe	Value of Output (Rs.)	Land Sown (Acres)	Yield	Fertilizer (Rs.)	Pesticide (Rs.)	Irrigation Exp. (Rs.)
Estimate	-1.070**	2662.105	-6.011	34.202	195.222	98.016	12.097
	(0.533)	(6409.768)	(11.500)	(450.035)	(680.059)	(222.821)	(98.590)
Observations	566	566	566	566	566	566	566
Elections	54	54	54	54	54	54	54

Note: Panel A shows that being in the winning faction does not change the farmer's access to government services. The second column shows that winners seek no more services than losers. The third and fifth colums show that winners borrow no more money and are no more likely to get loans from the government. The fourth column shows that conditional on taking loans the winners pay no less in interest. Panel B shows that being in the winning faction does not affect production. The data on production come form the 2006 and 1999 rounds of the survey, which typically fall within the current and previous village council. I must discard the other elections. The first column verifies that the sample is not too small to detect factionalism; winners are still less likely to pay a bribe. But the other columns show there is no evidence that winners produce more output or use more inputs.

REFERENCES

ACEMOGLU, D. AND J. ROBINSON (2012): *Why Nations Fail: Origins of Power, Poverty and Prosperity*, Crown Publishers (Random House).

ACEMOGLU, D. AND T. VERDIER (2000): "The Choice between Market Failures and Corruption," *American economic review*, 194–211.

ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2006): "Structural Identification of Production Functions," .

ADHVARYU, A., N. KALA, AND A. NYSHADHAM (2013): "Booms, Busts, and Household Enterprise: Evidence from Coffee Farmers in Tanzania," .

ALFARO, L. AND A. CHARI (2012): "Deregulation, Misallocation, and Size: Evidence from India," .

ALFARO, L., A. CHARLTON, AND F. KANCZUK (2008): "Plant-Size Distribution and Cross-Country Income Differences," .

ANDERSON, T. AND C. HSIAO (1981): "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 598–606.

——— (1982): "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of econometrics*, 18, 47–82.

BAKER, C. AND P. PHONGPAICHIT (2009): *A History of Thailand*, Cambridge University Press.

BANDYOPADHYAY, S. AND E. SKOUFIAS (2012): "Rainfall Variability, Occupational Choice, and Welfare in Rural Bangladesh," *World Bank Policy Research Working Paper*.

BANERJEE, A. AND E. DUFLO (2005): "Growth Theory Through the Lens of Development Economics," *Handbook of Economic Growth*, 1, 473–552.

——— (2007): "The Economic Lives of the Poor," *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 21, 141.

BANERJEE, A. AND B. MOLL (2010): "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2, 189–206.

BANERJEE, A. AND K. MUNSHI (2004): "How Efficiently Is Capital Allocated? Evidence from the Knitted Garment Industry in Tirupur," *Review of Economic Studies*, 71, 19–42.

BANERJEE, A. V. (1997): "A Theory of Misgovernance," *The Quarterly Journal of Economics*, 112, 1289–1332.

BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2009): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," .

BENJAMIN, D. (1992): "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models," *Econometrica*, 60, 287–322.

BESLEY, T., R. PANDE, L. RAHMAN, AND V. RAO (2004): "The Politics of Public Good Provision: Evidence from Indian Local governments," *Journal of the European Economic Association*, 2, 416–426.

BLISS, C. J. AND N. H. STERN (1982): *Palanpur: The Economy of an Indian Village*, Oxford University Press.

BOLLARD, A., P. J. KLENOW, AND G. SHARMA (2012): "India's Mysterious Manufacturing Miracle," *Review of Economic Dynamics*.

BUERA, F., J. KABOSKI, AND Y. SHIN (2011): "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101, 1964–2002.

DE MEL, S., D. MCKENZIE, AND C. WOODRUFF (2008): "Returns to Capital in Microenterprises: Evidence From a Field Experiment," *The Quarterly Journal of Economics*, 123, 1329–1372.

ENGLE, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica: Journal of the Econometric Society*, 987–1007.

FISMAN, R. (2001): "Estimating the Value of Political Connections," *American Economic Review*, 1095–1102.

GANDHI, A., S. NAVARRO, AND D. RIVERS (2011): "On the Identification of Production Functions: How Heterogeneous Is Productivity?" .

GOVT. OF KARNATAKA (1993): "Karnataka Panchayati Raj Act," .

GUMMERT, M. AND J. RICKMAN (2011): "When to Harvest," Tech. rep., International Rice Research Institute.

HOUSE, C. L. AND M. D. SHAPIRO (2008): "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *The American Economic Review*, 737–768.

HSIEH, C. AND P. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

IMBENS, G. W. AND T. LEMIEUX (2008): "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142, 615–635.

JENSEN, R. AND N. MILLER (2010): "A Revealed Preference Approach to Measuring Hunger and Undernutrition," Tech. rep., National Bureau of Economic Research.

JEONG, H. AND R. TOWNSEND (2007): "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, 32, 179–221.

JONES, C. I. (2011): "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, 3, 1–28.

KABOSKI, J. AND R. TOWNSEND (2009): "The Impacts of Credit on Village Economies," *MIT Department of Economics Working Paper No. 09-13*.

——— (2011): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 79, 1357–1406.

KARLAN, D., R. D. OSEI, I. OSEI-AKOTO, AND C. UDRY (2012): "Agricultural Decisions after Relaxing Credit and Risk Constraints," Tech. rep., National Bureau of Economic Research.

LEVINSOHN, J. AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of economic studies*, 70, 317–341.

MIDRIGAN, V. AND D. XU (2013): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*.

MIGUEL, E. AND M. K. GUGERTY (2005): "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 89, 2325–2368.

MOLL, B. (2010): "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *University of Chicago mimeo*.

MORDUCH, J. (1990): "Risk, Production and Saving: Theory and Evidence from Indian Households," *Harvard University, Manuscript*.

MUNSHI, K. AND M. ROSENZWEIG (2009): "Why Is Mobility in India So Low? Social Insurance, Inequality, and Growth," Tech. rep., National Bureau of Economic Research.

MURPHY, K. AND R. TOPEL (2002): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 20, 88–97.

OSOTIMEHIN, S. (2011): "Aggregate Productivity and the Allocation of Resources over the Business Cycle," .

PAGAN, A. (1984): "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25, 221–247.

PETERS, M. (2011): "Heterogeneous Mark-Ups and Endogenous Misallocation," *Manuscript, Massachusetts Institute of Technology*.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707–720.

ROSENZWEIG, M. AND H. BINSWANGER (1993): "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 103, 56–78.

ROSENZWEIG, M. R. (1988): "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries," *The Economic Journal*, 98, 1148–1170.

ROSENZWEIG, M. R. AND O. STARK (1989): "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *The Journal of Political Economy*, 905–926.

SHARMA, S. S. (1988): *Factions in Northern Rural India*, Shree Publishing House.

SHENOY, A. (2014a): "Market Failures and Misallocation: Separating the Costs of Factor and Financial Market Failures," .

——— (2014b): "Risky Income or Lumpy Investments? Testing Two Theories of Under-Specialization," .

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20.

SUNDARAM, R. (1996): *A First Course in Optimization Theory*, Cambridge Univ Pr.

TOWNSEND, R. (1994): "Risk and Insurance in Village India," *Econometrica: Journal of the Econometric Society*, 539–591.

TOWNSEND, R., A. PAULSON, S. SAKUNTASATHIEN, T. J. LEE, AND M. BINFORD (1997): "Questionnaire Design and Data Collection for NICHD Grant *Risk, Insurance and the Family* and NSF Grants," .

TOWNSEND, R. M. (2012): "Townsend Thai Project Monthly Survey (1-60) Initial Release," Murray Research Archive [Distributor] V1 [Version].

UDRY, C. (1994): "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria," *The Review of Economic Studies*, 61, 495–526.

——— (1996): "Gender, Agricultural Production, and the Theory of the Household," *Journal of Political Economy*, 104, 1010–1046.

WALKER, T. S. AND J. G. RYAN (1990): *Village and Household Economics in India's Semi-Arid Tropics*, Johns Hopkins University Press.

WOOLDRIDGE, J. (2002): *Econometric Analysis of Cross Section and Panel Data*, The MIT press.

YANG, D. AND H. CHOI (2007): "Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines," *The World Bank Economic Review*, 21, 219–248.