NBER WORKING PAPER SERIES

EXTERNAL VALIDITY IN A STOCHASTIC WORLD

Mark Rosenzweig
Christopher Udry

Working Paper 22449
http://www.nber.org/papers/w22449

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2016

The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w22449.ack

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Mark Rosenzweig and Christopher Udry. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

External Validity in a Stochastic World
Mark Rosenzweig and Christopher Udry
NBER Working Paper No. 22449
July 2016, Revised August 2016
JEL No. C93,O1,O13,O14,O15

ABSTRACT

We examine the generalizability of internally valid estimates of causal effects in a fixed population over time when that population is subject to aggregate shocks. This temporal external validity is shown to depend upon the distribution of the aggregate shocks and the interaction between these shocks and the casual effects. We show that returns to investment in agriculture, small and medium enterprises and human capital differ significantly from year to year. We also show how returns to investments interact with specific aggregate shocks, and estimate the parameters of the distributions of these shocks. We show how to use these estimates to appropriately widen estimated confidence intervals to account for aggregate shocks.

Mark Rosenzweig
Department of Economics
Yale University
Box 208269
New Haven, CT 06520
and NBER
mark.rosenzweig@yale.edu

Christopher Udry
Economic Growth Center
Yale University
Box 208269
New Haven, CT 06520
and NBER
udry@yale.edu

I. INTRODUCTION

A large fraction of empirical work in economics is concerned with estimating causal effects. These causal estimates are often used as the basis for policy recommendations and sometimes even policy initiatives. As empirical methods improve, most markedly with the rapid adoption of RCTs, it is appropriate to shift our focus of concern from issues of internal validity (e.g., identification) towards the external validity of estimates of causal effects. What can we learn from internally valid estimates of a causal effect of a treatment or action in a particular population in a specific set of circumstances about the effect of that same treatment or action in other populations and circumstances? Recently there have been studies that have examined how a particular estimated relationship varies by the population sampled (Allcott, 2015) or across countries (Dehejia et al., 2016). These show that generalization of treatment effects across populations can be difficult: in the first case because of the potential for systematic selection bias in the populations that are studied; in the second because of limited information about the underlying heterogeneity across populations.[1] Banerjee et al. (2014, 2016) provide a general framework for considering the external validity of causal estimates, arguing that researchers should explicitly provide "structured speculation" regarding the scope and degree of external validity of their estimates. We implement this suggestion and develop an approach that specifies the heterogeneity that is relevant for generalizing causal estimates in one population at one time to other populations or circumstances, and a model of the interactions between those dimensions of heterogeneity and the causal effect under consideration.

We examine an important dimension of external validity that has been relatively ignored. Our focus recognizes that the world is stochastic and subject to aggregate shocks - weather shocks, technological shocks, price shocks – that can affect the entire population simultaneously. Therefore, for the same population the aggregate environment may vary from year to year. To what extent is an internally valid causal estimate obtained at a point in time externally valid for the same population at a different point in time? This is the question of *temporal* external validity.

If causal estimates are state-dependent it is possible that they have little external validity even for the same population. This is because almost all RCTs and most other empirical investigations carry out their intervention or explore a relationship for a single period, or, if over multiple periods, do not pay attention to whether the causal effects are heterogeneous over time.[2] Our understanding of external validity across populations also depends on the degree of temporal external validity: for example, the comparison across countries in Dehejja et al. (2016) uses data based largely on single-year estimates of the causal effect. The variation observed across countries may be partially due to a failure of temporal external validity.[3]

---

[1] Dehajia et al. (2016) provide a useful review of other recent work on external validity.

[2] A small set of examples (that are either influential or by one of us) includes Duflo et al. (2011), Banerjee and Duflo (2008), Banerjee et al. (2013), Bloom et al. (2013), de Mel et al. (2008, 2009), Foster and Rosenzweig (1995), Hanna et al. (2014), Karlan et al. (2013), Fafchamps et al. (2011), Udry and Anagol (2006), Suri (2011).

[3] The cross-sectional variation in optimal agricultural decisions discussed in Hana et al. (2014) may also be a consequence of treatment effects that vary over time – a lack of temporal external validity.

An important implication of an acknowledgement that causal effects might vary with the realization of aggregate shocks is that the confidence interval (CI) of the estimated parameters from a single-year study does not provide guidance for external validity in a stochastic setting. If aggregate shocks vary and such shocks alter the causal effects, the true confidence interval should also incorporate the intertemporal distribution of the external shocks and will be wider than the estimated CI. The true confidence interval is typically the decision-relevant measure, both for the agents and for policymakers.

There have been recent examinations of the temporal external validity of two influential studies. Justin Lin's estimate (AER, 1992) of the effect of agricultural reforms in China was shown by Zhang and Carter (AJAE, 1997) to be too large. Lin (1992) did not consider the effects of weather on agricultural output. In particular, the estimates of the effect of agricultural reforms presented in Lin (1992) relied on data from only one post-reform year. That year had considerably more favorable rainfall than the one pre-reform year included in the analysis. Zhang and Carter estimate that improvements in weather accounted for 8% of the post-reform growth in output examined by Lin (high rainfall)

Ironically, the deworming intervention at the heart of Miguel and Kremer (2004) also occurred during exceptional weather – the first year was the hottest year recorded in Kenyan history, and rainfall was at historically high levels. Zhang (2016) shows that as a consequence, helminth infection rates in the study years were more than double the average rate over the 35-year period 1978-2012. Deworming treatment is almost 100% effective in removing hookworm, roundworm and schistosomiasis infections, so the health impact of the treatment is increasing in the initial rate of infection. As a consequence, it is likely that the causal impact of the deworming intervention on health estimated in Miguel and Kremer (2004) overstates the expected impact in a different year.[4] Zhang (2016) uses cross-sectional variation in initial infection rates to estimate the effect of the realization of this aggregate shock on the treatment effect of deworming on later infection rates and concludes that the estimated treatment effect is approximately double the effect that would be realized in a more typical year. In section 4, we consider in more detail the potential for using cross-sectional variation in the realization of large scale shocks to assess temporal external validity.

The empirical issues we address are whether aggregate shocks affect an important set of casual relationships – the returns to investment in agriculture, enterprises and schooling – and if so how important is the variation in the aggregate shocks. We use four different panel data sets to measure the extent of variability in investment returns over time and to examine the determinants of intertemporal investment return variability. In the specific contexts of these datasets, we identify some of the particular aggregate shocks that affect investment returns. We focus on returns to agricultural investment, to investment in non-farm small and medium enterprises, and to investment in human capital because these are critical parameters for diagnosing the barriers to economic development and thus are essential for policy.

[4] However, the probability of reinfection will be higher in an environment with high background infection rates, which may mitigate this bias.

We also show how it is possible to estimate the true confidence interval for any single estimate, when one can characterize the distribution of the stochastic shocks and have an estimate of the sensitivity of the causal impact to the realization of the shocks. In a stochastic world the investment decisions of agents, whether entrepreneurs, individuals or policy-makers typically depends upon the distribution of these shocks. Clearly, no single year estimate of the returns given a particular realization of the shocks is sufficient to characterize this distribution.

In section 2 we set out a simple model of investment decisions of agents in a stochastic environment in which the return to investment is subject to aggregate shocks. We provide a formula for combining information on the distribution of the aggregate shocks and estimates of the effect of the shocks on marginal returns to investment that generates confidence intervals around the expected return to investment that incorporate the uncertainty generated by the macro shocks.

In order to generate decision-relevant confidence bounds, we require estimates of the sensitivity of investment returns to particular aggregate shocks and the parameters describing the distribution of the shocks. In sections 3 and 4 we estimate these parameters in the context of agriculture using panel data on individual farmers and rainfall outcomes in India and Ghana. In both settings, rainfall is highly variable and as a consequence, agricultural profits are highly variable over time. The average over farmers of the intertemporal coefficient of variation (cv) of a farmer's profits in India is 1.5 and in Ghana is 0.9. We show that the returns to planting stage investments in both contexts is very sensitive to rainfall realizations and that rainfall realizations are themselves quite variable. We show, as a consequence, that the probability that a single year estimate of the rate of return to agricultural investment is within a reasonable bound of the expected return is very low. For example, the probability that a single year estimate of the rate of return is within 30 percentage points (on either side) of the expected value of the rate of return is only 5% in India and 28% in Ghana.

To show that the sensitivity of returns to investment to aggregate shocks is not limited to agriculture, in section 5 we examine the returns to investment in small enterprises in Sri Lanka, using panel data on microenterprises collected by De Mel et al (2008). As we found in agriculture, there is significant variation over time in the profits realized by each entrepreneur (cv=0.5), which is consistent with respondents' subjective expectation of profit variability. We estimate a strong seasonal pattern in the returns to investment by these entrepreneurs net of individual illness shocks. This pattern is consistent with respondents' reports that cite aggregate demand fluctuations as an important source of profit variation.

We cannot identify the aggregate shocks driving variation in returns to investment in Sri Lanka. In section 6, we examine specific macro factors that drive inter-annual fluctuations in the returns to schooling in Indonesia. An object of interest in many studies in developed and developing countries is the rate of return to schooling. The most influential studies recognize the existence of individual-specific unobservables that jointly affect schooling and earnings, just as for agricultural and enterprise investments, and attempt to estimate the returns to schooling exploiting policy variation relevant to schooling attainment choices as instruments (e.g., Angrist and Krueger, 1991; Angrist and Krueger, 1992;

Card and Lemieux, 2001; Card, 1995; Duflo, 2001). These studies too, however, ignore stochastic variation in earnings returns associated with aggregate shocks. For example, Angrist and Krueger (1991) use variation in the timing of births combined with variation in compulsory schooling laws in the United States as instruments determining schooling and then look at earnings outcomes separately from the 1970 and 1980 Censuses. The differences between the estimated returns to schooling across the two Census years are comparable to or much larger than (in the most complete specification) the difference between the IV and OLS estimates within the same year that is the focus of the study.[5] Duflo (2001) is one of the few studies based on data from a developing country that exploits program variation to identify schooling returns, but she uses earnings from one Census year to obtain her estimates. We will examine in more detail below to what extent, if any, stochastic variation in macro shocks limits the external validity of her estimates.

There is evidence that aggregate factors have persistent effects on schooling returns. For example, macro shocks, measured by rates of unemployment occurring at the time of labor market entry evidently affect the returns to schooling, since such shocks differentially affect the initial earnings of college and high-school graduates (Oreopoulos *et* al., 2012). There is also a literature that attempts to identify the sources of longer-term trends in the returns to schooling, but sorting out the causal effects of longer-term economy-wide supply and demand factors has proved difficult.[6]

We examine short-term changes in the return to schooling. We use panel data on urban wage and salary workers over the period 1993-2008, during which there were dramatic changes in the world price of oil and a rapid depreciation of the Rupiah as a consequence of the 1998 financial crisis. We combine data on these shocks with information on industry-specific skill intensity, openness to trade, and the distribution of industries across provinces to estimate the effects of these shocks on the returns to schooling in each year. We find that these external macroeconomic shocks had important effects on the return to schooling as estimated in any year, in line with our expectations of how these shocks affect the relative demand for skilled and unskilled workers. These findings are used to evaluate the temporal external validity of Duflo's (2001) estimate of the rate of return to education in Indonesia, based on earnings data from 1995. This ex-post analysis of temporal external validity can serve as an example of how researchers can provide "structured speculation" regarding the likely degree of temporal external validity of the causal estimates of their work.

In the conclusion, we provide recommendations for the use of theory and external data to improve, or at a minimum assess, the temporal external validity of causal estimates.

---

[5]In 1970, the IV point estimate of the return to schooling is 10.1% (se=.033); in 1980 it is 7.0% (se=.03). There is no discussion of why the returns across years differ but it is interesting to note that in the reference earnings year in the 1970 Census (1969) the unemployment rate was 3.5%, while in 1979 the unemployment rate was 5.8%.

[6]For example, Krueger (1993) estimates how changes in computer use in the workplace affected the returns to schooling between 1984 and 1989 in the US. DiPietro and Pedace (2008) examine shifts in skilled employment by industry as factors influencing the demand for schooling in Argentina from 1995 to 2003.

II. FRAMEWORK

We consider a population of agents who take actions in period zero and realize returns in period one. In a stochastic world in which this population may be subject to aggregate shocks, what can be learned about the expected return or the distribution of returns to those actions from the realized returns in a particular state of nature? This is a problem of *external validity*; while the population may be identical, the realized returns to a particular action may vary across these different potential states.

To fix ideas, agent *i* takes an action (a_0) now with a current cost and a future benefit:

$$\max_{a \in A} c_{i0}(a_0, z_{i0}, s_0) + E[V_{i1}(a_0, s_1, \zeta_i)|\Omega_0] \tag{1}$$

The source of uncertainty in the model is the random vector s_1. z_{i0} is a characteristic of i that might affect that agent's costs; ζ_i is a permanent characteristic that might influence the future benefit. Ω_0 is the agent's information set at the time of the choice of a_0. We will define $a_0 \in \mathbb{R}$ so that $c_{i0}(a_0)$ is decreasing and convex in a_0. s_0 is a variable that might or might not affect costs of the action in period zero, and is something that is known by the agent at the time the action is chosen that might contain information informative of the distribution of s_1.

The agent's expected future return is

$$E[V_{i1}(a_{0i}, s_1, \zeta_i)|\Omega_0] = \int V_{i1}(a_{0i}, s_1, \zeta_i) f(s_1|\Omega_0) ds_1$$

While s may be multidimensional, we will typically restrict our attention to a single dimension, and to simplify notation we define

$$\beta_i(a, s_1, \zeta) \equiv \frac{\partial V_{1i}(a, s_1, \zeta)}{\partial a}.$$

To an analyst with goal of estimating the return to the action, z is a candidate for an instrumental variable, or for an intervention that might induce a change in a. This will be the case in three of the examples below.

Three aspects of the random vector s_1 are particularly important for our analysis. First, its variability. A single estimate of the average rate of return to an action is more informative in an environment in which the variability of s_1 is more limited. Second, the responsiveness to realizations of s_1 of the marginal value of increasing a_0. For given variability of s_1, the larger is this responsiveness, the less informative a single estimate of the average rate of return. Finally, its dependence on the initial signal s_0. As s_0 becomes more informative, both the responsiveness of agents' choices to s_0, and the usefulness for policy of estimates of the dependence of returns on realizations of s_1 increase.

Suppose that s_0 contains information about s_1 ($f(s_1|\hat{s}_0) \neq f(s_1|\tilde{s}_0)$ for some $\hat{s}_0 \neq \tilde{s}_0$, so. It will now typically be the case that the optimal choice of a depends on the realization of s_0, if $\frac{\partial \beta}{\partial s_1} \neq 0$. For example, if increases in s_0 cause a first order stochastic dominant shift in the distribution of s_1. Then

$$\frac{\partial E\beta_i(a^*, s, \zeta^*)}{\partial s_0} = \int \beta_i(a, s_1, \zeta) \frac{\partial f(s_1|s_0)}{\partial s_0} ds_1.$$

which is positive (negative) if $\frac{\partial \beta_i}{\partial s_1} > 0\ (< 0)$. The signal influences the optimal choice of action:

$$\frac{da_{i0}}{ds_0} = -\frac{\frac{\partial E\beta_i(a, s_1, \zeta)}{\partial s_0} + \frac{\partial c_{i0}}{\partial s_0}}{\frac{\partial^2 c_{i0}}{\partial a_0^2} + \frac{\partial^2 E\beta_i}{\partial a_0}}.$$

The denominator is negative, so if there is no direct effect of s_0 on the cost of action a_0 (as, for example, with a weather forecast), then

$$sign\left(\frac{da_{i0}}{ds_0}\right) = sign\left(\frac{\partial \beta_i}{\partial s_1}\right).$$

As we would expect, an informative signal that the realization of the aggregate shock is likely to be larger leads to greater (less) investment if the shock and investment are complements (substitutes). Direct effects of the signal on costs of the action (perhaps via market prices of policy responses to the signal) may modify this response. However, conditional on the agent's choice and the realization of the state s_1, s_0 is irrelevant for the return to the investment. This makes s_0 a second attractive candidate for an instrumental variable.

In a typical experiment or econometric exercise, the primary object of interest is the impact of a on an outcome Y. Y might be the agent's welfare Vat differing levels of a: $V_{i1}(a = 1, s_1) - V_{i1}(a = 0, s_1)$ or in the case of a continuous action, $\frac{\partial V_{i1}}{\partial a}$. More typically, Y is an observable outcome other than V, such as the monetary profit generated by the investment $\pi_{i1} = r_i(s_1, \zeta_i)a$. The standard evaluation problem is that, of course, we can never simultaneously observe different levels of *a* for the same individual *i.* That is not our focus in this paper, so we will assume that the usual concerns with the internal validity of our estimate have been solved. Therefore, we assume $V_{i1}(a, s_1, \zeta_i) = V_1(a, s_1, \zeta_i)$ and when we narrow our focus to the monetary rates of return to investment, we specify $V_1(a, s_1, \zeta_i) = V_1(\pi_{i1}) = V_1(r(s_1, \zeta_i)a)$.

We have data on $\{a_{i0}, z_{i0}, s_0, Y_{i1}, \zeta_i, s_1\}_{i \in I}$. Y_{i1} can be V_{i1}, π_{i1}, or any other observable outcome of the investment decision that might (or might not) depend upon the realization of the aggregate shock. In the empirical work, we focus on monetary rates of return, so $Y_{i1} = Y_1(a, s_1, \zeta_i) = r_i(s_1, \zeta_i)a$. In line with our assumption that internal validity is unproblematic, we suppose that ζ is fully observed, that z or s_0 are appropriate instruments or that a_{i0} can simply be randomized with perfect compliance. Indeed, we begin by assuming that our sample is sufficiently large and our methods sufficiently robust that we have in hand

$$\hat{\beta}(a, \hat{s}_1, \zeta) = \beta(a, \hat{s}_1, \zeta) \equiv \frac{\partial Y_1(a, \hat{s}_1, \zeta)}{\partial a} \tag{2}$$

over the range $a \in A$ for $s_1 = \hat{s}_1$ and for all observed ζ.

Even in this benign scenario, in which our confidence interval around the estimated $\hat{\beta}(a^*, \hat{s}_1, \zeta^*)$ has collapsed, it remains the case that we observe this return function at only a single value of the aggregate shock s_1. This information may be of relatively limited value, however, because it is typically the expected impact $E(\beta)$ that is relevant for policy decisions. The tight confidence interval we estimate

conditional on the realization of the aggregate state necessarily understates the breadth of the confidence interval that would cover the policy relevant value $E(\beta)$.

In some contexts, it may be possible to calculate explicit bounds on the likelihood of any particular estimate of the returns to an action being close to the expected value of that return. Consider a unidimensional continuous shock. We observe $\beta(a, s_1, \zeta)$, asymptotically without error, for $\{a, \zeta\} \in A \times$ Z, but only for $s_1 = \hat{s}_1$. For any particularly value of $a = a^*,\ \zeta = \zeta^*$, how likely is it that our estimate of the return to *a* is within δ of $E\beta(a^*, s_1, \zeta^*)$? To simplify notation, define

$$\frac{\partial^2 Y_1(a^*, s, \zeta^*)}{\partial a \partial s} = \frac{\partial \beta(a^*, s, \zeta^*)}{\partial s} = \alpha$$

When Y is defined as the profit generated by the investment, as in the empirical exercise below, $\alpha = \frac{\partial^2 \pi_1(a^*, s, \zeta^*)}{\partial a \partial s} = \frac{\partial r(s, \zeta^*)}{\partial s}$.

Then a first order approximation implies

$$\begin{aligned} prob[(E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) - \delta) \leq \beta(a^*, s_1, \zeta^*) \leq (E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) + \delta)] \\ \approx prob\big[(E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) - \delta) \leq E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) + \alpha\big(s_1 - E(s_1|\hat{s}_0)\big) \\ \leq (E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) + \delta)\big]. \end{aligned}$$

Therefore,

$$\begin{aligned} prob[(E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) - \delta) \leq \beta(a^*, s_1, \zeta^*) \leq (E(\beta(a^*, s_1, \zeta^*)|\hat{s}_0) + \delta)] \\ \approx prob\left[-\frac{\delta}{\alpha} \leq \big(s_1 - E(s_1|\hat{s}_0)\big) \leq \frac{\delta}{\alpha}\right] \end{aligned} \tag{3}$$

For any given value of δ, a larger standard deviation of *s* or a larger cross derivative $\frac{\partial^2 Y_1(a^*, s, \zeta^*)}{\partial a \partial s}$ reduces the probability that a given observation of the return to an action is near the expected value of that return. If s is normally distributed, then the probability of estimating a return to *a* that is within $\alpha\sigma_s$ of its expected value is approximately 68%.

For k-dimensional s, (3) generalizes with

$$\boldsymbol{\alpha}(a^*, \boldsymbol{s}_1, \zeta^*)|\hat{s}_0) \equiv D_{s_1}\beta(a, \boldsymbol{s_1}, \zeta)$$

to

$$\begin{aligned} prob[(E(\beta(a^*, \boldsymbol{s}_1, \zeta^*)|\hat{s}_0) - \delta) \leq \beta(a^*, \boldsymbol{s}_1, \zeta^*) \leq (E(\beta(a^*, \boldsymbol{s}_1, \zeta^*)|\hat{s}_0) + \delta)] \\ \approx prob\big[-\delta \leq \boldsymbol{\alpha}'\big(\boldsymbol{s}_1 - E(\boldsymbol{s}_1|\hat{s}_0)\big) \leq \delta\big] \end{aligned} \tag{4}$$

These calculations require information on α (or $\boldsymbol{\alpha}$). Here, we will estimate α in three different settings, to assess its importance and to calculate confidence bounds that incorporate the distribution of relevant macro shocks.

III. ICRISAT VILLAGE SURVEYS 2005-2011

We use panel data from the ICRISAT Village Dynamics in South Asia (VDSA) surveys for the years 2005-2011 to estimate the returns to planting-stage investments and their sensitivity to rainfall realizations. The data are based on surveys of 30 farmers in each of the six villages from the first generation ICRISAT VLS (1975-1984). The villages are located in the states of Maharashtra (4) and Andhra Pradesh (2). The ICRISAT data set contains farmer-level investment and profit data for seven consecutive years, permitting us to quantify the sensitivity of investment returns to aggregate shocks. There are three additional features of these data relevant to our investigation. First, input and output information is provided in approximately three-week intervals collected by resident investigators. This enables us to precisely measure investments made within a season prior to the realization of rainfall (s_1 in the model) as well as the season-specific profits associated with those investments. Second, there are data on daily rainfall for each of the six villages for as long as 26 years. This enables us to both estimate the influence of rainfall realizations on investment returns and to characterize the distribution of rainfall states $f(s)$ faced by farmers so that we can compute confidence intervals that take into account stochastic outcomes. Rainfall and thus profit variability in the ICRISAT villages is high.[7] Figure 1 displays mean profits by year for the four villages from which we obtain our estimates from 2005-2011. The average coefficient of variation in profits over the period experienced by the ICRISAT farmers is 1.4.

The third feature of the ICRISAT data that we exploit is that for the four villages in Maharashtra, the forecasts of monsoon rainfall issued by the India Meteorological Department (IMD) in late June have been moderately successful in predicting *kharif*-season (July-September) rainfall and, most importantly, significantly influence the planting-stage decisions of the ICRISAT farmers (Rosenzweig and Udry, 2014) in the July-August planting stage. We use the IMD forecast for the southern peninsula as an instrument in an IV strategy for estimating the returns to planting-stage investments (the value of labor used in plowing, seeding and fertilizing plus the costs of the material inputs) and their sensitivity to rainfall variation when such investments are endogenous. In the context of the model, the IMD forecast is a signal to farmers (s_0) that affects their expectations of seasonal rainfall $f(s_1|s_0)$ and thus profitability but, for given investments, rainfall realization and prices, should have no direct effect on profitability. In particular, we estimate a conditional profit function using the ICRISAT panel data, treating planting-stage investments as an endogenous choice that responds to the rainfall forecast. To identify investment returns, we have to impose additional structure to ensure that the forecast instruments satisfy the exclusion restriction.

Agricultural profits depend on investments in planting-stage inputs and on the realization of rainfall, and as our model has emphasized, on the interaction between these. In addition, agricultural profits are functions of a number of dimensions of heterogeneity, such as farm size, soil characteristics, and irrigation and interactions of these with rainfall. There is also good evidence (Sharma and Acharya 2000) that profits depend as well on lagged rainfall (differentially depending upon farm characteristics, particularly soil depth) through the soil moisture overhang effect. Hence we specify a linearized version

[7] Our measure of profits is the value of agricultural output minus the value of all agricultural inputs, including the value of family labor and other owned input services.

of the farm profits of household h in village v in year t that is quadratic in planting-stage investments that includes a farmer fixed effect and village-year fixed effects that absorb time-varying village-specific input prices (particularly wages) that could be correlated with rainfall forecasts. A key feature of our specification is that it allows the effects of planting-stage investments on profits to depend on the realization of rainfall. Excluded from the profit specification are the rainfall forecast and its interactions with exogenous fixed land characteristics. This is the primary identification assumption required to estimate the returns to planting-stage investments. That is, conditional on realized rainfall (or village-year fixed effects) the forecast of total rainfall in the monsoon affects profits only through its effect on investments. There are two primary concerns regarding this excludability assumption. The first is that conditional on our specific measures of realized rainfall, the forecast of total rainfall may be correlated with an unmeasured dimension of rainfall that matters for profits. We measure realized rainfall as the total amount of rainfall over the year and the total amount of rainfall over the monsoon, as the IMD long-range forecast is the prediction for the total amount of rainfall over the monsoon. Binswanger and Rosenzweig (1993) have shown that the monsoon onset date is a salient feature of rainfall for farm profits in India. However, in the ICRISAT data we find that conditional on even a subset of our measures of rainfall (monsoon rainfall), the IMD forecast of total monsoon rainfall is not correlated with the onset date. Note that the village-year effects capture all time-varying aggregate shocks.

Second, the rainfall forecast for a given year is common to everyone in a village. Through its effect on input demand, a forecast of good (bad) weather could raise (lower) input prices - particularly wages - in a village. In principle it is possible as well that there could be policy interventions (changes in regulated grain prices, emergency agricultural interventions, *ex ante* efforts to provide relief). These village-specific changes correlated with the forecast could affect profits directly. As noted, the village-year fixed effects are included in the profit function to address this possibility. A casualty of including village-year fixed effects is that the direct effects of rainfall and lagged rainfall on profits are not identified.[8]

[8]A further concern that would make the forecast non-excludable is that the increased planting-stage investments induced by a favorable forecast reduce the farmer's resources available for subsequent production stages. In the model this is ruled out by the implicit assumption of perfect credit markets within the relevant production cycle. The ICRISAT survey data enable us to carry out a global separability test similar to that of Benjamin (1992). The basic idea is that exogenous changes in the family labor force should not affect profits if all input markets are unconstrained. Illness has a large random component (net of the household fixed effect), and illness can affect the family's ability to supply labor. For the years 2005, 2006, 2010 and 2011 the ICRISAT survey elicited information on the number of days that adult family members were ill in the *kharif* season. Household fixed effect estimates obtained for the total sample of farmers and the farm households in the Maharashtra villages of the effect of the number of sick days on total labor days in the kharif season indicate that for each day an adult was sick almost a third of a day of on-farm family labor was lost. The estimate is $LabDays_{it} = \underset{(.09)}{-0.34} Sickdays_{it} + \underset{(.0004)}{.002} \cdot TotRain_{vt}$. If liquidity constraints limited the ability of the household to substitute hired labor to make up for family labor days lost, an increase in sick days should therefore decrease profits. However, FE-IV estimates of the profit function for the Maharashtra farmers including the number of adult sick days (not reported) indicate that we cannot reject the hypothesis of separability - despite sick days evidently significantly reducing on-farm family labor supply, an increase in the number of adult sick days has no impact on profitability.

Table 1 reports fixed-effects instrumental variable (FE-IV) estimates of the profit function, with the FE at the farmer and village-year levels. The IMD forecast interacted with the characteristics of the farm and farmer are the instruments for planting-stage investments. All profit function specifications include the rainfall variables interacted with total landholdings, irrigated landholdings, soil depth, and four soil types (red, black, sandy, loam). The first column of Table 1 reports estimates from the profit specification that is linear and quadratic in investment. Based on those estimates we can strongly reject the hypothesis that larger planting-stage investments do not increase profits over almost the full range of the investment distribution in the sample. In the second column we add rainfall interactions with the investment variables. We can also reject that investment returns do not depend on rainfall. These estimates thus imply that *ex-post* optimal investments depend on realized rainfall outcomes, or, put differently, how much under-investment one would infer from profit function estimates depends on what is assumed to be the typical rainfall outcome. The estimates in column 2 imply that at mean levels of investment in the ICRISAT sample, returns to planting-stage investment are positive over the full range of rainfall realizations observed in the data, as shown in Figure 2.

The point-wise confidence intervals depicted in Figure 2 for the rainfall-specific returns to agricultural investment assume that the only source of variation in the estimates is sampling error. However, our estimates suggest that variability in rainfall also strongly affect returns. The true confidence interval for any estimate obtained for a given rainfall realization s_t should reflect as well the variability in the distribution of s, that is, $f(s)$. To construct the true confidence interval, and to assess the probability that any one estimate is within pre-specified bounds around expected profitability (profit at mean rainfall) we use the rainfall time series from Kinkheda, the ICRISAT village with the longest continuous history - 26 years. Figure 3 displays the distribution of annual rainfall from the village over the 26-year period. Standard tests of normality indicate non-rejection of the null (e.g., Shapiro-Wilk W test [p=.28]). We thus use the standard distribution and mean of the rainfall distribution to assess the true confidence interval. We do this first assuming that we know α exactly, using our estimate of alpha from column 2 of Table 1. Based on the actual rainfall distribution parameters and α, we draw 1000 profit returns estimates. The distribution of profit returns, taking into account rainfall variability, is shown by the line in Figure 4.

We also computed the distribution of returns based on both the standard deviation of our estimates (due to sampling variability) and rainfall variability, assuming that the two errors are independent. That distribution is depicted by the cross-hatched line in Figure 4. We contrast these with the distribution of estimates implied by just sampling variability, reported in all studies, for two returns estimates, at the 25th and 75th percentile of the rainfall distribution. As can be seen, confidence intervals based on sampling variation alone severely understate the true confidence interval constructed from the distribution of estimates incorporating both sampling variation and stochastic variability in rainfall.

Based on the estimated distribution of returns, we can compute, based on (3), the probabilities that any one estimate of returns obtained in a random, single year lies within some bounds around the true expected investment return (at mean rainfall). The first column of Table 2 reports these probabilities, for intervals ranging from 10 percentage points to 50 percentage points on either side of the mean. For example, the estimates indicate that the probability that a single-year's estimate of profit returns in the

ICRISAT village lies within 10 percentage points on either side of the expected return is less than 2 percent; and there is only a 9% chance of even being within 50 percentage points on either side of the expected return for one-year study. Note that these probabilities represent best cases, as they do not incorporate sampling error and because they are solely based on rainfall variability, which is unlikely to be the only source of aggregate shocks that affect demand.[9] The external validity of an estimate of an investment return obtained in a single year is evidently extremely low in the ICRISAT setting.

IV. RETURNS TO AGRICULTURAL INVESTMENT IN NORTHERN GHANA, 2009-2012

We use three years of panel data from 1352 households in 75 communities in the Northern Region of Ghana to examine the variability in the returns to planting stage investments as a function of weather realizations over these three years across this region. The large number of communities and their broad spatial extent raises the possibility of using cross-sectiional variation in weather realizations to estimate α; we use this data to compare cross-sectional and fixed effect estimates of the responsiveness of returns to weather shocks. This data set contains detailed plot level information on inputs at each stage of production and seasonal output, with geographical information on the location of each plot that can be combined with data on weather realizations over the season. An important advantage of these data is that they were collected in the context of a randomized controlled trial which varied the availability of rainfall index insurance and substantial grants of cash across the sample of farmers and over time.[10] Karlan et al. (2014) show that the availability of rainfall index insurance to farmers in the sample generates a strong investment response. Farmers who are insured with rainfall index insurance (as a consequence of randomly being provided provided grants of free or subsidized insurance) increase their investment in cultivation by an average of $266 (s.e. $134) over a baseline expenditure of approximately $2058. The same RCT showed that the allocation of a farmer's investment in cultivation was influenced by the receipt of a cash grant: expenditure on fertilizer increased by $56 (s.e. $17) upon receipt before the cultivation season began of a cash grant averaging $420. The randomized allocation of this sample into a variety of subsamples facing exogenously varying budget constraints provides us with exogenous variation in planting season investments which we use to estimate the return to these investments across varying realization of weather.

As in the ICRISAT villages, there is a great deal of variability in rainfall and thus profits in northern Ghana. The average coefficient of variation in net output (the value of harvest minus the cost of all purchased inputs) for each farmer over the three years is 0.9. The fact that Karlan et al. find strong responses of investment to access to rainfall index insurance provides *a priori* evidence that investment returns vary by rainfall. They find in addition that the treatment effect in a given year of having access to

[9] To the extent that rainfall variability is correlated with other shocks that impinge on profits (e.g. pests, temperature), we cannot say that we have identified that part of returns variation due solely to rainfall.

[10] See Karlan et al (2014), section III and online appendix 1 for a detailed description of the sample, data collection procedures, index insurance and cash grants interventions and the randomization.

insurance on harvest values is higher for households in areas that receive higher rainfall in that year. Karlan et al., however, do not directly investigate the effect of weather realizations on the rate of return to investment, nor do they explore the implications of the sensitivity of investment returns to weather for the distribution of expected returns.

We take advantage of the index designed for rainfall insurance to provide a second dimension, along with total rainfall, to characterize the state of nature (s_{1vt}).[11] This insurance product provides an index (I_{vt}) of the weather realization which combines information on the amounts and timing of daily rainfall to predict harvest for the most important single crop in the region, which is maize. The index is constructed from daily rainfall data available over the period 1983 – 2013 that we obtained from the Ghana Agricultural Insurance Pool (who have developed and market the successor rainfall index insurance product).[12] We combine this index with information on total rainfall over the growing season (R_{vt}) to create a two-dimensional indicator of the weather shock in each community in each year $s_{1vt} = (R_{vt}, I_{vt})$.

Net plot income depends on planting season investments, on the realization of s_{1vt}, and on their interaction. In Ghana we focus on net income from the plot, defined as the value of all output from the plot minus the cost of purchased inputs and hired labour because of the difficulty of measuring and valuing the use of family labor on the plot. Hired labor tends to be used for specialized tasks and at peak moments of labor demand and therefore the measured wage overstates the opportunity cost of household labor. The planting season investments we examine include clearing, field preparation and fertilizer application. We use the random assignment of households to varying treatments of cash grants and grants of or subsidies to rainfall index insurance, and interactions of these treatments with baseline plot area and land characteristics as instruments for plot level investment. Identification therefore relies on the assumption that the assignment to alternative treatments affects net income only through the choice of planting season investments. The primary concern that arises with respect to this assumption is that conditional on planting season investments, assignment to the different treatments could influence later-stage cultivation decisions and thus be correlated with net income. This will not be the case for assignment to the cash grant treatment if liquidity constraints do not bind with respect to expenditures within the growing season. The results of Karlan et al. (2014) showing that these households were able to substantially increase average planting season investments upon assignment to free or reduced cost insurance without any infusion of additional capital suggest that within season liquidity constraints are not binding, as we found in section 3 for the ICRISAT sample. The identification assumption with respect to the insurance treatments is that conditional on the level of planting season investment and rainfall realizations, post-planting cultivation decisions are independent of insurance. Identification is threatened, however, if there is sufficient flexibility in cultivation opportunities after the conclusion of planting season investment that farmer decisions might be influenced by insurance status (e.g., conditional on planting

[11] In contrast to the significantly more dry ICRISAT India context, in northern Ghana excessive rainfall is a concern of farmers, and their dependence on maize increases the sensitivity of yields to the timing of rainfall during the season.
[12] We verify the accuracy of these data for the second half of this period using the Tropical Rainfall Measurement Mission at a 0.25 x 0 .25 degree resolution "The TRMM and Other Data Precipitation Data Set, TRMM 3B42," Huffman and Bolvin (2015)

stage cultivation decisions and rainfall realizations, an insured farmer decides to replant after a late season drought spell while an uninsured farmer does not). This identification concern is mitigated by the fact that by the completion of planting season, farmers have accrued 90% of total non-family labor costs for the entire season.[13]

We begin by examining the variation in the returns to planting season investment over time and across space in northern Ghana. We divide the 75 communities in the sample into 10 clusters of geographically proximate communities, based on their proximity to a TRMM grid point and estimate the returns to planting season investment separately for each cluster-year, with no effort to associate any variation in the return to investment with weather realizations in that cluster-year. The results are reported in Figure 5. The community clusters are identified by color, and there is a separate estimate of the return to investment for each year and each cluster. The black lines show 95% confidence intervals around each estimate (the three cluster-year estimates for which no confidence interval is visible have small standard errors). It is immediately apparent that the estimated return to planting season investment varies dramatically both across clusters and within clusters over time. To take a typical example, community cluster 9 has an estimated return to planting season investment of -36% in the first year, -15% in the second, and +15% in the third year of the survey. The magnitude of the variation in returns to planting season investment over time within a cluster appears to be approximately the same as the variation over clusters within a year. The mean across clusters of the standard deviation of returns over time within each cluster is 0.54, while the standard deviation across clusters of the mean (over time) of returns within each cluster is 0.39.

What drives this dramatic variation in estimated returns to investment across these cluster-years? We examine the influence of weather realizations on realized returns to planting season investments in Table 3. Table 3 reports the results from household fixed effects instrumental variable estimates of the net income function. The first column shows that net income depends on planting stage investments, but in contrast to the results in the ICRISAT villages, the average return to planting stage investment is negative at the mean level of investment (*r*=-0.68 (s.e. 0.15)).[14] In column 2, we show interactions between planting stage investments and the vector of weather realizations. These interactions are jointly significantly different from zero ($\chi^2(6) = 19.63\ p = 0.00$) and imply that the returns to planting stage investments at the sample mean weather realization and level of investment are approximately zero (net rate of return = -0.39, s.e. 0,14).

The estimates in column 2 of Table 3 are used to calculate the returns to planting stage investments over the range of the rainfall realization index I_{vt} in the data set, conditional on total rainfall being held at its median. These estimates, presented in Figure 6 of the rate of return range from -42% in

[13] Estimates from a narrower definition of "planting season" that excluded fertilizer expenses are available from the authors (and were provided in an earlier version of the paper). The results are qualitatively similar to those presented here. Absolute returns are higher, and as would be expected if there is flexibility in cultivation decisions in response to rainfall realizations (Fafchamps, 1993) in the early season, the returns to the more narrow definition of the planting season investment are more variable than those we present here, with the more expansive definition of the planting season.

[14] The table provides estimates of gross returns, (1+r).

the worst case rainfall realization to +40% in the best case. The vertical red line is drawn at the median weather realization of the full distribution of outcomes over the 1983-2013 period and indicates that expected returns are .08 at that median outcome. The point -wise confidence intervals presented in Figure 6 are based on the assumption that sampling error is the only source of uncertainty in the estimate of the return to investment. However, our estimates imply a very strong response of net returns to the realization of weather in any season. While these confidence bounds may be accurate estimates of the statistical uncertainty surrounding our estimate of the rate of return conditional on a particular weather realization, they (perhaps strongly) underestimate the level of uncertainty regarding the expected value of the rate of return. The appropriate confidence interval around any particular estimate of the rate of return to this investment must reflect as well the underlying variation in the distribution ofs_{vt}. In order to investigate the implications of this variability, we use the 31 years of rainfall data for each community to estimate the joint distribution of (R_{vt}, I_{vt}).[15] We then use draws from this estimated distribution to examine the variability in the realized returns to investment that will be generated by the variability in weather conditions in northern Ghana.

Our primary goal is to understand the degree to which observation of the return to agricultural investment in a particular season conditional on a particular aggregate weather realization provides information about the expected value of this return. Therefore, we begin by assuming that our experiment-induced variation in investment is pristine and that our sample size is sufficiently large that we observe the rate of return precisely given any specific weather realization. Further, let us begin by assuming that the variation in weather across communities and over the three years of the Ghana sample has provided us with sufficient information to estimate the dependence of these rates of return on weather realizations (the vector $\boldsymbol{\alpha}$) with similar precision. Based on our estimates of the rainfall distribution and $\boldsymbol{\alpha}$, we can simulate the distribution of realized returns to investment. This is reported as the solid line in Figure 7. Of course, we only estimate $\widehat{\boldsymbol{\alpha}}$, so Figure 7 also reports the distribution of expected returns taking into account both rainfall variability and the sampling error in our estimates $\widehat{\boldsymbol{\alpha}}$.

In contrast, Figure 7 also reports the distribution of expected returns generated by sampling error alone at two specific realizations of $s_1 = (R, I)$ – "good rain" that generates an expected net return equal to the 75th percentile of the overall distribution of net returns, and "poor rain" that generates an expected net return equal to the 25th percentile of the overall distribution. The confidence interval one would calculate from the standard errors of the net income function at either of these two specific realizations would dramatically understate the width of the confidence interval constructed from the distribution of estimates that incorporates both sampling variation and the underlying variability in weather conditions.

Given our estimates of the distribution of weather realizations and $\widehat{\boldsymbol{\alpha}}$, we can use (4) to calculate the probability that any single estimate of the returns obtained in a random season drawn from this

[15] We cannot reject normality of the distribution of R_{vt} for any of the communities in our data. There is no evidence of serial correlation in total rainfall in the savanna zone of Ghana (Manzanas et al., 2014), so we parameterize total rainfall in any community as a draw from a normal distribution with a mean and standard deviation equal to our sample values of those parameters. The weather index I_{vt} takes on eight values, and we estimate the probability of realization of each of these values separately for each of the four quartiles of the overall rain distribution.

rainfall distribution lies within some distance δ of the true $ex\ ante$ expected return to investment. These probabilities are reported in the second column of Table 2, for $\delta \in (10, 20, \ldots 50)$ percentage points. The probability that an estimate of the returns to investment in agriculture in northern Ghana drawn from a single season in a single location is within 10 percentage points of the expected return in northern Ghana is 9%. We have only a 34% chance of estimating a return that is within even 40 percentage points on either side of the true expected return to investment in any sample containing a single draw of weather. Weather variability is less extreme in northern Ghana than in the Indian ICRISAT villages, but the external validity of an estimate of investment returns is extremely low for agriculture in northern Ghana.

The results in Table 2 present a significant challenge to researchers interested in the returns to investment, new technologies, or market innovations in agriculture. An alternative to estimating a from multiple observations of a single population over time could be to exploit cross-sectional variation in s_1 over space at a given time. Using a cross section would require abandoning the fixed effects specification. This would only be appropriate if the realization of s_1 is uncorrerlated with fixed unobservable characteristics that might affect net profits. The data in northern Ghana encompass a sufficiently broad geographic area (150 km across its greatest extent) and rainfall realizations are sufficiently locallized that there is significant intra-annual variation in weather realizations across communities. Therefore, it is possible to estimate the returns to planting stage investment and the interactions of those returns with weather realizations in a single cross-section. The results from estimating the relationship between rainfall realizations and returns to investment (evaluated at the median level of investment and median total rainfall) using each of the 3 waves of the panel are compared to those using the full panel and fixed effects in Figure 8. The full panel results are identical to those presented in Figure 7. As shown in the Figure, however, estimated returns from the cross sections vary widely across the three years, and exhibit a substantially different relationship between rainfall realizations and returns than is generated by the fixed effect estimates.[16] This is not surprising given that we expect the cross-sectional variation in the realization of weather to be correlated with variation in characteristics of the distribution of weather, which in turn will in general be related to returns to investment.

V. SRI LANKA MICRO ENTERPRISES 2005-2007

De Mel *at al.* (2009) carried out a panel survey of 408 micro-enterprises (less than $1,000 in invested capital) in three southern and southwestern districts of Sri Lanka for nine consecutive quarters between 2005 and 2007. The aim of the study was to estimate the returns to investment by implementing an RCT in which cash and capital equipment were offered to a randomized subset of firms after the first survey round and after the third round. We principally use their data to explore whether the issue of returns variability due to external aggregate shocks is salient outside of the agricultural sector. The authors of the study pooled the quarterly profit data and provided one estimate of capital returns for the

[16] We reject the equality across cross-sectional years of the set of coefficients that generate these curves. The test statistic is distributed as F(17,7206), and is 3.71 (p=0.000). The coefficients tested are those on planting stage investment, its square and its interactions with weather realizations and plot area (akin to the specification reported in Table 3, column 2).

whole survey period. They believed that the profit variability they observed was predominantly measurement error, and indeed trimmed the top .5% of quarter-to-quarter differences. We use their trimmed data and their IV strategy, which exploited the randomized interventions as instruments for investments, to estimate investment returns by quarter.

A component of the survey instrument sheds light on the issue of whether the measured profit variability in the data is due to external aggregate demand shocks or was simply measurement error. The entrepreneurs were asked to provide information on their expectations about their enterprise's profits in the subsequent quarter. Specifically, they were asked to indicate the lowest amount of monthly profits they thought they could earn three months from the survey date and the highest. The answers indicated that the respondents anticipated significant real quarterly profit variability. We computed the triangular coefficient of variation based on the answers, which was 0.5. Interestingly, the average coefficient of variation across firms for measured quarterly profits over the nine quarters was also 0.5.

More interestingly, the entrepreneurs were also asked to name the most important source of variability in profits. While 42.9% gave illness as the most important reason (own or family), 32% indicated that the most important factor causing a deviation in profits was an aggregate shock to product demand, including weather variation (8.2%), changes in income among customers (14.0%), and disruptions of infrastructure (road closing, power outages), 10.3%.

The first column of Table 4 reports our enterprise and quarter fixed effects IV estimates of the return on investment from the enterprise data that pools the profit data across quarters, as in the original study. These estimates are obtained from the same unbalanced panel of 385 firms with at least three rounds of information as used by de Mel *at al.* Our specification differs from theirs in two ways, however. First, we allow the investment responses to the treatments to vary by year and by the gender of the entrepreneur, as in subsequent work the same authors found that treatment response was significantly greater for male respondents (de Mel *et al.*, 2012). Second, given the reports by the respondents on the importance of illness in influencing profits, in order to isolate demand and supply shocks we include a dummy variable indicating whether or not the entrepreneur reported any illness that resulted in his or her losing days of work. The estimate of illness variable coefficient is statistically significant, and indicates that when the entrepreneur is ill, profits are lower by 10%. The return on investment is also statistically significant, using the conventional standard error as our criterion, and is similar in magnitude to that found in the original study.

In the second column we report estimates from a specification that allows the returns to vary by quarter. In all quarters returns are positive and statistically different from zero. The point estimates, however, vary from .066 to .133, a factor of 100%. The difference between the lowest quarterly return and the highest is statistically significant. However, we cannot reject the hypothesis by conventional standards that the returns are equal across all quarters.

Because of the redundancy of quarters across years - for example, the April-June quarter is sampled twice - we can test for the existence of a seasonal pattern in returns based on the quarterly returns estimates. We cannot reject the hypothesis of a seasonal pattern (similarity of returns across seasons). We thus obtained new estimates of investments returns that are permitted to vary by season rather than by quarter. These estimates are reported in the first column of Table 5. For this specification, all the seasonal investment returns are significantly different from zero, but we can reject the hypothesis that the returns do not differ across seasons. There is real seasonal returns variability. And the point estimates indicate that the difference between the highest and lowest seasonal returns is a statistically significant 33%. Moreover, the pooled estimate of the returns to investments, net of season and entrepreneur fixed effects (column 2), .098, is not equal to the mean of the seasonal estimates of .087.

VI. RETURNS TO SCHOOLING: URBAN INDONESIA 1993-2000

In this section we examine whether aggregate shocks affect schooling decisions and the extent to which they affect the realized rate of return to education using panel data on urban wage workers in Indonesia. To fix ideas about how shocks affect the returns to schooling and the relationship between single-year estimates of returns and the expected value of returns, we set out a simple two-period Mincer model of schooling in which agents choose whether to attend secondary school and occupation-specific wages are subject to time-varying shocks. Thus, in terms of our general model $A = \{0,1\}$ denoting attaining secondary education or not and

$$\begin{aligned} c(a, z, s_0) &\equiv w_l(s_0) - a(w_l(s_0) + z) \\ V(a, s_1, \zeta) &\equiv \gamma(1-a)w_l(s_1) + a\gamma w_h(s_1, \zeta), \end{aligned} \tag{5}$$

where z is the direct cost of schooling (say, school access costs), γ is the discount rate, w_l is the wage of labor in unskilled occupations (A=0) and w_h is the wage of labor in skilled occupations where laborers must have A=1. As before, s_0 is the first-period shock, which may have predictive power for the second-period shock s_1. ζ is a characteristic of the skilled occupation that affects the way in which the realized shock affects the skilled wage.

Ignoring individual heterogeneity in productivity differences across occupations, an agent will choose to attend secondary school if

$$w_l(s_0) + \gamma E_s w_l(s_1) < \gamma E_s w_h(s_1) - z$$

Thus, changes in z, changes in the first period shock s_0, and changes in expected wage difference across occupations will affect schooling decisions. Note that if the first period shock s_0 affects wage expectations then it will affect schooling decisions by both affecting the opportunity cost of schooling and the expected gains from schooling. However, the effect of variation of z, given the linearity of the cost function, is independent of the realization of s_0.

Given perfect mobility across occupations, and stationarity in shocks, across many cohorts the expected or average rate of return to schooling will be equal to the inverse of the discount rate:

$$\frac{1}{\gamma} = \frac{E_s\big(w_h(s_1,\zeta) - w_l(s_1)\big)}{E_s(w_l(s_0)) + z},$$

from which it can be seen that in the long run, while changes in the direct cost of schooling will affect schooling decisions and thus how many people choose secondary schooling there will be no change in the return to schooling when there is a permanent change in schooling costs z. Thus when school costs are reduced, there will be a decline in the skilled wage relative to the unskilled wage (because more people attend secondary school), but no change in the rate of return if the lower school costs are considered in calculating the return. There will be no general-equilibrium effect on the return.

While schooling choices are determined by the expected value of the return to education at the time of the schooling decision, what is typically estimated is the "return" to schooling in one year

$$\frac{w_h(\hat{s}_1,\zeta) - w_l(\hat{s}_1)}{w_l(s_0)}.$$

If the shock in that year differentially affects skilled and unskilled wages, then the return estimated in that year may be quite different from the expected return that rational agents use to make schooling decisions.[17]

How the shock affects the return depends on the source of the shock. We will consider as examples of shocks an increase in the price of oil or a change in the exchange rate, and consider the schooling decisions of families across Indonesia's provinces. The distribution of these macroeconomic shocks is the same across all provinces, but they may have varying impacts on the distribution of expected wages and thus schooling decisions. For exchange rate shocks we denote the industrial structure of a province by ζ as the share of high skilled employment in industries that are more open to trade. In either case, imperfect factor mobility across industries and provinces and transportation costs within Indonesia imply that

$$\frac{\partial^2 w_h}{\partial\zeta\partial s} > 0.$$

The wage premium for schooling is increased more in provinces with a higher share of high skilled employment in open to trade industries when the exchange rate depreciates. The expected value of the wage premium is equalized (at $\frac{1}{\gamma}$) across provinces, but the slope of the relationship between the realized wage premium in a single year and these macroeconomic shocks varies according to the industrial structure of the province.

To assess to what extent and how macro shocks affect estimates of the returns to schooling we

[17] Jensen (2010), for example, presents a single year estimate of the skilled-unskilled wage differential to households making schooling decisions to test whether lack of information on schooling returns is a factor explaining low schooling. How the respondents responded to the estimate would depend on their knowledge of how informative a single year's estimate of the schooling return was for gaging the expected return, which would in turn depend on the realization of the shock in that year relative to its expected value.

use panel data from the first three rounds of the Indonesia Family Life Survey (IFLS), focusing on short-term contemporaneous variation in macro shocks, as we did for agricultural and non-agricultural investments. We selected the IFLS, a panel survey based on a representative sample of households in 13 major provinces in Indonesia begun in 1993, for three reasons: First, we can construct an annual hourly wage series for individual (wage and salary) workers from 1993-2000. The survey rounds were in 1993, 1997, and 2000, but every wave has a retrospective earnings history so we can fill in the inter-survey years using the wage and employment histories.

A second reason for selecting the IFLS for this time period is that Indonesia was affected by two external macro shocks. Indonesia is a major producer and exporter of petrochemicals and thus the Indonesian economy is potentially significantly affected by intertemporal variation in the world price of oil, which experienced considerable variability in the relevant time period. In addition, the financial crisis of 1998 substantially altered the value of Indonesian currency on the world market, potentially affecting the workforce in industries engaged in international trade relative to those in industries producing non-tradeables. Figure 9 shows the movements in the price of oil and the rupiah exchange rate from 1990 to 2000, both of which display considerable volatility over the period. We can assess to what extent these macro shocks affected the returns to schooling in urban areas in the eight years of the panel.[18]

The third reason for selecting the IFLS data is that, with our estimates of the sensitivity of schooling returns to oil price and rupiah variation in the decade of the 1990's, we can assess the external validity of Duflo's estimate of the returns to schooling, which was based on a school-building initiative in the 1973-78 period and earnings measured from the 1995 Indonesia census.

In our panel sample of male urban wage and salary workers aged 16-80 with eight years of non-zero earnings from the IFLS, the average coefficient of variation in earnings is a nontrivial .43. This is comparable to our estimate of micro-enterprise profit variation from the Sri Lanka panel. Some of this is variation is also due to measurement error. In these data we can obtain an upper bound on measurement error by comparing the overlap of contemporaneous wage reports and retrospective wage reports from a subsequent survey round (for example, in the 1997 round, respondents report their wage in 1993, the preceding survey year). We find that the measurement error in wage rates accounts for approximately 40% of the cross-sectional variance, assuming independence of errors. But, it is unlikely that there is as much measurement error in earnings as there is in agricultural or micro-enterprise profits computed from survey data, where we found evidence of systematic factors that altered investment returns. The main issue is whether the returns to schooling vary significantly across years, and to what extent this variation is attributable to aggregate shocks.

The first column of Table 6 reports the OLS estimate of "the" rate of return to schooling pooling

[18]There have been studies of the effects of oil price shocks on the US economy. These have shown that from 1948 to 1972, increases in the world price of oil played a significant role in US recessions (Hamilton, 1983) but had no detectable adverse effects after 1973 (Hooker, 1996). These results are not obviously applicable to Indonesia, as the petrochemical industry constitutes a much smaller share of industry employment in the United States than in Indonesia. And governmental revenues in the US are not nearly as dependent on oil revenues.

all observations for urban male wage and salary workers with at least two years of non-zero earnings over the eight-year panel. The specification is the standard "Mincer' log wage specification, including schooling years, age and age squared but we also include year fixed effects. The set of individual linear year coefficients are statistically significant, and indicate significant variation in hourly wage levels across years. In the next column, the rates of return are allowed to vary by year. All of the year-specific returns are statistically different from zero by conventional standards, and we can also reject the hypothesis that all the year-specific returns are equal.

The OLS returns estimates in the first two columns may be biased due to the endogeneity of schooling. In the third column of Table 6 we report worker fixed effects estimates, which eliminates the influence of any worker characteristics that jointly affect schooling and earnings. In this specification we cannot identify the returns to schooling, but we can obtain consistent estimates of differences in returns across years. The year-specific returns coefficients in the third column are relative to that in the omitted year, 2000. These estimates too lead to rejection of the hypothesis that the schooling returns are equal across years. The point estimates indicate, for example, that the schooling returns in 1998 and 1999 are from 2.6 to 2.7 percentage points, respectively, below that for 2000. These differences are statistically significant

Just as for the agricultural samples, we are interested in understanding the sources of variation in returns. We not only want to estimate α's for schooling returns but also to explore the mechanisms. We thus examine the influence on the return to schooling investment of variation in the world price of oil over the sample period and of changes in the rupiah exchange rate in the last three years of the panel, when we think the variation in the value of the rupiah is mainly due to forces external to Indonesia rather than to domestic factors. With respect to the oil price, as noted, whether an increase in the price of oil increases or decreases the schooling return in the short run depends on whether it raises the demand for skilled relative to unskilled labor. We would expect that an increase in the oil price would likely raise the returns to schooling. First, an increase in the price of oil would lead to an expansion in the demand for workers in the petrochemical industry, and the 1990 Census indicates that workers in the petrochemical industry are more skilled relative to the overall workforce, whether measured by mean years of schooling or proportion of workers with secondary education. Second, because government revenues depend importantly on the price of oil in Indonesia, we would expect that employment in the public sector would also expand, and generally public sector employees are more skilled. For example, the INPRES school-building program in 1973-78, financed with the new revenues resulting from an increase in the price of oil in that period, increased as well the employment of teachers, who are significantly more educated than the average Indonesian worker. We thus would expect that in the short run, before workers can re-allocate among industries, a rise in the price of oil will benefit skilled relative to less skilled workers and raise the return to schooling.

We re-estimate the log wage equation omitting the set of individual year-schooling interactions and replacing them with an interaction between the oil price and schooling to estimate α for oil price variation. Because the oil price is time-varying, we can identify the influence of shocks to oil prices on

schooling returns while including the worker fixed effect.[19] Column 1 of Table 7 reports the OLS estimates of the return to schooling and of α. The specification includes schooling, age, age squared, the interactions of schooling and the age variables with the oil price, the oil price, and year. As expected α is positive and statistically significant, while the direct effect of the price of oil on wages (for the unskilled) is actually negative. When the worker fixed effect is included in column 2, the estimate of the oil price α retains its statistical significance, and increases in magnitude by 50%, while the direct oil price effect is slightly attenuated.

The point estimate of α in the second column implies that a one standard deviation (measured over the decade) increase in the price of oil would modestly raise the return to schooling, by a statistically significant 0.7 percentage points. However, the estimate also implies that the rise in the price of oil from 1998 to 2000 ($16 per barrel) increased the return to schooling by 2.5 percentage points, accounting for all of the observed increase in the return to schooling in that interval seen in column two of Table 7.

Our model and estimate of the oil price α allow us to *ex post* evaluate the external validity of Duflo's estimates of both the effects of increasing access to schools via school building on schooling attainment and the returns to schooling (Duflo, 2001). In that study, the massive program building schools in Indonesia in the years 1973 through 1978 was used to estimate the effects of increasing school access on schooling and the consequences of the increased schooling for earnings, based on the earnings for the relevant cohorts in 1994-1995.

Figure 10 displays the changes in oil prices over the period 1960 through 2016, with the price bars for the relevant program and earnings years relevant to the Duflo study darkened. What is apparent is that the beginning of the school building initiative was marked by a permanent increase in oil prices. This is not surprising, as the financing of the program was supported by the increase in oil revenues experienced by the government. The jump in oil prices at the time of the program corresponds to s_0 in the model, and if that increase affected expectations of future higher levels of oil prices, given our estimate of a positive α, the oil price increase alone would have increased schooling attainment by altering the expected rate of return to schooling even if access to schools had remained unchanged (that is, had z remained constant). A comparison of schooling attainment before and after the program (single difference) would thus have likely over-estimated the effect of the program on increasing enrollment. The methodology used, however, not only differenced cohorts schooled before and after the program but also across areas with different program intensity. This double difference, given linearity in the cost function as in our model, identifies the program effect on schooling attainment that is independent of the influence of the contemporaneous oil price shock, whatever its effect on expectations. The estimate of the impact of the school building program on schooling attainment (for Indonesia) thus has internal validity as well as full external validity with respect to stochastic (oil price) shocks.

The estimate of the rate of return to schooling, however, was obtained based on earnings

[19] 2.7% of respondents moved to another province over the eight-year period. These respondents were dropped. In the fixed effects specification, only if such movers had schooling returns unusually sensitive to the oil price shocks will this affect (attenuate) our estimates of the shocks on schooling returns.

measured in a single year, from the 1995 Census. An issue is what is the appropriate return to compare Duflo's single-year estimate? One benchmark might be the expected or average return over the decade, say the return associated with the decade average oil price. The average oil price from 1991-2000 was \$18.2, while the price of oil in the Census earnings reference year used by Duflo was between \$15.5 and \$16.9, a difference of only about \$2 from the mean price. Given our estimate of α, this difference implies that Duflo's estimate of the expected return over the decade was too low by only 0.3 percentage points - 1995 just happened to be a typical year. However, if Duflo had used 2000 as the reference year, say from the 2000 Indonesian Census, when the oil price had spiked to \$27.6, our estimate of α implies that her estimate of the schooling return would have been too high by about 1.5 percentage points, 21% above the expected value for the decade, given her externally valid estimate of the schooling return of 7%.

We have assumed that the oil price increased the relative demand for skilled versus unskilled workers in Indonesia as the reason why we found the oil price α to be positive. Here we more directly test the hypothesis that changes in schooling returns due to macro shocks arise from the effects they have on the relative demand for skilled and unskilled labor, exploiting the movements in the exchange rate value of the rupiah over the 1997-200 period. As seen in Figure 9, the exchange rate increased threefold between 1997 and 1998 due to the worldwide financial crisis in 1998. To carry out the analysis, we make use of a method employed in the trade literature to examine the effects of changes in trade policies on schooling choice. The key idea is that the policies (e.g., tariff reforms) affect the returns to schooling via their effects on industries differentiated by skill intensity and sensitivity to the policy - those industries open to trade (e.g., Edmonds *et al.*, 2010; Atkin, 2015). In this literature, the estimates are of the effects of the shifts in trade variables via industry structure on schooling choice. Here we estimate the effects of the external trade shock on schooling returns, which influence the decisions of agents making schooling choices. Analogous to the inquiries in the trade literature, we wish to test whether in areas (provinces) where there are more workers employed in industries that are both more open, and thus disproportionally affected by a change in the value of the rupiah, and more skilled, an exogenous rise in the exchange rate raises the returns to schooling. That is, the more open are skilled industries in an area the greater the increase in the return to schooling.

To carry out this analysis we constructed a standard measure of trade openness by industry ((exports + imports)/ value added) O_k , where k indexes industry. We obtained industry-specific information on imports and exports in 1990 for Indonesian industries from the UN ComTrade database and industry output and province- and industry-specific GDP share for each province j G_{jk} from Statistics Indonesia.[20] We also obtained information on the skill-intensity (share of workers with secondary schooling) of industries S_k from the 1990 Indonesia Census. From these data we created our average skill/openness measure by province: $SO_j = \sum_k O_k S_k G_{jk}$. We then interacted that measure by the year-specific rupiah exchange rate xr_t and the schooling of the respondent. The hypothesis we are testing is that the coefficient on the triple interaction between (provincial) openness, (annual) exchange rates and (individual) education ($SO_j \cdot xr_t \cdot a_{0i}$) should be positive - a rise in the exchange rate increases the returns to schooling more in areas in which there is a heavier concentration of industries that are open and skill-

[20] http://comtrade.un.org/db/default.aspx and https://www.bps.go.id/linkTabelStatis/view/id/1200.

intensive.

The last two columns of Table 7 report the OLS and fixed effects estimates of the log wage equation, based on the years 1997-2000, that add to the specification including the oil price variables the exchange rate and openness/skill variables. As can be seen the returns to schooling do rise when the exchange rate increases in areas with more open and skilled industries. Increases in the oil price also increase the schooling rate of return in this sub-period. The point estimate from the individual fixed-effect specification indicates that at the mean of the *SO* measure of openness/skill for Indonesia the large rise in the exchange rate from its value in 1997 to its average value in 1998-2000 after the financial crisis increased the rate of return to schooling by 2.4 percentage points. If we again use Duflo's (externally valid) estimate of the mean return level for the decade, that is an increase of over a third in the schooling return. Clearly, domestic schooling returns are sensitive to global macro shocks in Indonesia. While exchange rate volatility in Indonesia after 2000 seems to be relatively low, as can be seen in Figure 10 oil price volatility in the first 15 years of the 21st century was substantially higher than it was in the 1990's. We would expect that therefore investments in schooling have become much more risky.

VII. CONCLUSION

We have examined to what extent causal estimates of the impact of decisions, whether about polices, investments, or technology choices, obtained in one time period have external validity within the same population but under different circumstances due to the existence of time-varying aggregate shocks. We showed that the key set of parameters needed to assess temporal external validity is that describing how aggregate shocks affect the returns on the actions, α. Estimates of shock effects on returns, however, are rarely obtained in existing studies, even in those studies using information from multiple years, which could in principle be used to identify α. We also showed that to calculate appropriate estimate confidence intervals where macro shocks are important requires not only information on the α but also knowledge of the parameters of the shock distributions.

In our empirical investigation of the sensitivity of estimates of investment returns to external macro shocks we have found evidence that investment returns, whether returns to agricultural investments, enterprise investments or investments in human capital, differ significantly from year to year. And, where we could measure the impact of specific macro shocks, such as for agriculture and for schooling, we found that the α's were large in magnitude and in sign consistent with economic theory. Moreover, where we could also reliably estimate the parameters of the distribution of one key macro shock, rainfall in Ghana and India, we showed that the likelihood of a single year estimate being within reasonable bounds around the expected investment return are very low. Thus, for rain-fed agriculture, single-year estimates of the returns to investments or interventions have extremely low temporal external validity. Our estimates also imply that our knowledge of the impact of investments or policy interventions may be far less certain than existing estimates of confidence intervals imply. Agents making investment decisions in a stochastic world face substantially more risk than measured by conventional estimates that ignore aggregate shocks.

What is the remedy for enhancing temporal external validity and assessing the external validity of estimates in a world with stochastic shocks? First, micro models of individual decision-making need to identify and incorporate the role of aggregate shocks. Theory should suggest which states of nature matter, and how, for the outcomes and interventions of interest. Second, information on the time-series of the relevant states of nature (shocks) for the population or site being studied should be obtained along with information on the realization of the shocks for the study period. Ideally, identification of α should also become a component of any well-designed study. This likely requires information spanning multiple periods, as cross-sectional shock variation can be correlated with site-specific unobservables that influence outcomes. Generally, places with an adverse state of nature in a given year may be more likely to have bad outcomes, which will affect ex ante behavioral responses to interventions and outcomes. And we have shown using data from Ghana that the cross-sectional estimates of the effects of macro shocks do not accurately represent the effects of shocks for given fixed characteristics.

Typically, though, an analyst using data from one particular context may not have the data available to estimate α. Knowledge of the context, of agents' decision-making processes and relevant technologies, however, can generate a prior belief regarding its sign and magnitude. These, coupled with estimates from external sources of the distribution of the relevant aggregate shocks can guide informed speculation regarding the applicability of any particular estimate of the returns to actions to those from future actions.

REFERENCES

Allcott, Hunt. (2015). “Site Selection Bias in Program Evaluation.” *Quarterly Journal of Economics* 130(3): 1117-1165.

Angrist, Joshua and Alan B. Krueger. (1991). “Does Compulsory School Attendance Affect Schooling and Earnings?” *The Quarterly Journal of Economics*, 106(4), 979-1014.

Angrist, Joshua and Alan B. Krueger. (1992). “Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery.” NBER Working Paper No. 4067.

Atkin, David. (2015). “Endogenous Skill Acquisition and Export Manufacturing in Mexico.” NBER Working Paper No. 18266.

Banerjee, Abhijit, Sylvain Chassang, and Erik Snowberg. (2016). “Decision Theoretic Approaches to Experiment Design and External Validity.” NBER Working Paper No. 22167.

Banerjee, Abhijit, and Esther Duflo (2008). “Do Firms Want to Borrow More: Testing Credit Constraints Using a Targeted Lending Program,” BREAD Working Paper No. 005, 2004, revised 2008.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster and Cynthia Kinnan. (2015). “The miracle of microfinance? Evidence from a Randomized Evaluation,” *American Economic Journal: Applied Economics* 7(1): 22-53.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. (2013). "Does Management Matter? Evidence from India," *Quarterly Journal of Economics* 128: 1-51.

Cai, Hongbin, Yuyu Chen, Hanming Fang, and Li-An Zhou. (2015). “The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Natural Field Experiment.” *Review of Economic and Statistics* 97(2): 287-300.

Card, David. (1995).“Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” *Aspects of Labour Economics: Essays in Honour of John Vanderkamp*, Louis Christofides, E. Kenneth Grant and Robert Swindinsky, eds. University of Toronto Press.

Cole, Shawn, Xavier Giné, and James Vickery. (2013). “How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment.” Harvard Business School, World Bank, Federal Reserve Bank of New York.

Dehejia, Rajeev, Cristian Pop-Eleches, and Cyrus Samii. (2016). “From Local to Global: External Validity in a Fertility Natural Experiment.” NBER Working Paper No. 21459.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2008). "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 123(4), 1329-1372, November.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2009). “Measuring Microenterprise Profits: Must We Ask How the Sausage Is Made?” *Journal of Development Economics*, 88(1): 19-31.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2012). “One Time Transfers of Cash or Capital Have Long Lasting Effects on Microenterprises in Sri Lanka.” *Science* 335: 962-966.

Di Pietro, Georgio and Lucas Pedace. (2008). “Changes in the Returns to Education in Argentina.” *Journal of Applied Economics* 11(2): 259-279.

Duflo, Esther. (2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” American Economic Review 91(4): 795-813.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. (2011). "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 101(6): 2350–90.

Duflo, Esther, and Christopher Udry. (2004). “Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices.” National Bureau of Economic Research working paper No 10498.

Edmonds, Eric V., Nina Pavcnik and Petia Topalova. (2010). “Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform.” *American Economic Journal: Applied Economics* 2(4): 42-75.

Fafchamps, Marcel. (1993). “Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers.” Econometrica 61(5): 1173-1197.

Foster, Andrew D. and Mark R. Rosenzweig. (1995). “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture.” *Journal of Political Economy* 103(6): 1176-1209.

Giné, Xavier, and Dean Yang. (2009). “Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi.” *Journal of Development Economics* 89 (1) (May): 1–11. doi:10.1016/j.jdeveco.2008.09.007.

Hamilton, James D. (1983). “Oil and the Macroeconomy since World War II.” *Journal of Political Economy* 91(2): 228-248.

Hanna, Rema, Sendhil Mullainathan Joshua Schwartzstein. (2014). “Learning Through Noticing: Theory and Evidence from a Field Experiment.” *The Quarterly Journal of Economics*: 1–43.

Hooker, Mark A. (1996). “What happened to the oil price-macroeconomy relationship?” *Journal of Monetary Economics* 38:195-213.

Jensen, Robert. (2010). “The (Perceived) Returns to Education and the Demand for Schooling.” *The Quarterly Journal of Economics*. 515-548.

Karlan, Dean, Isaac Osei-Akoto, Robert Osei, and Chris Udry. (2014). “Agricultural Decisions after Relaxing Credit and Risk Constraints,” *Quarterly Journal of Economics*, 129(2): 597-652.

Krueger, Alan B. (1993).“How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989*.” The Quarterly Journal of Economics* 108(1): 33-60.

Lemieux, Thomas and David Card. (2001). “Education, Earnings, and the 'Canadian G.I. Bill’.” *The Canadian Journal of Economics* 34(2): 313-344.

Lin, Justin Y. (1992). “Rural Reforms and Agricultural Growth In China." *American Economic Review* 82:34-51.

Miguel, Edward and Michael Kremer. (2004). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities.” *Econometrica* 72(1): 159–217.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. (2012). “The Short- and Long-Term Career Effects of Graduating in a Recession.” *American Economic Journal: Applied Economics* 4(1): 1–29.

Suri, Tavneet. (2011). “Selection and Comaparative Advantage in Technology Adoption.” *Econometrica* 79(1): 159-209.

Udry, Christopher and Santosh Anagol. (2006). “The Return to Capital in Ghana.” American Economic Review Papers and Proceedings 96.2: 388-93.

Zhang, Bin and Colin A. Carter (1997). “Reforms, the Weather, and Productivity Growth in China's Grain Sector.” *American Journal of Agricultural Economics*: 1266-1277.

Zhang, Warren. 2016. “Worm Wars: Do Mass Deworming Campaigns Really Improve Education?” Manuscript: Yale University

Figure 1. Mean Real Profits in Four ICRISAT Villages 2005-2011, by Year

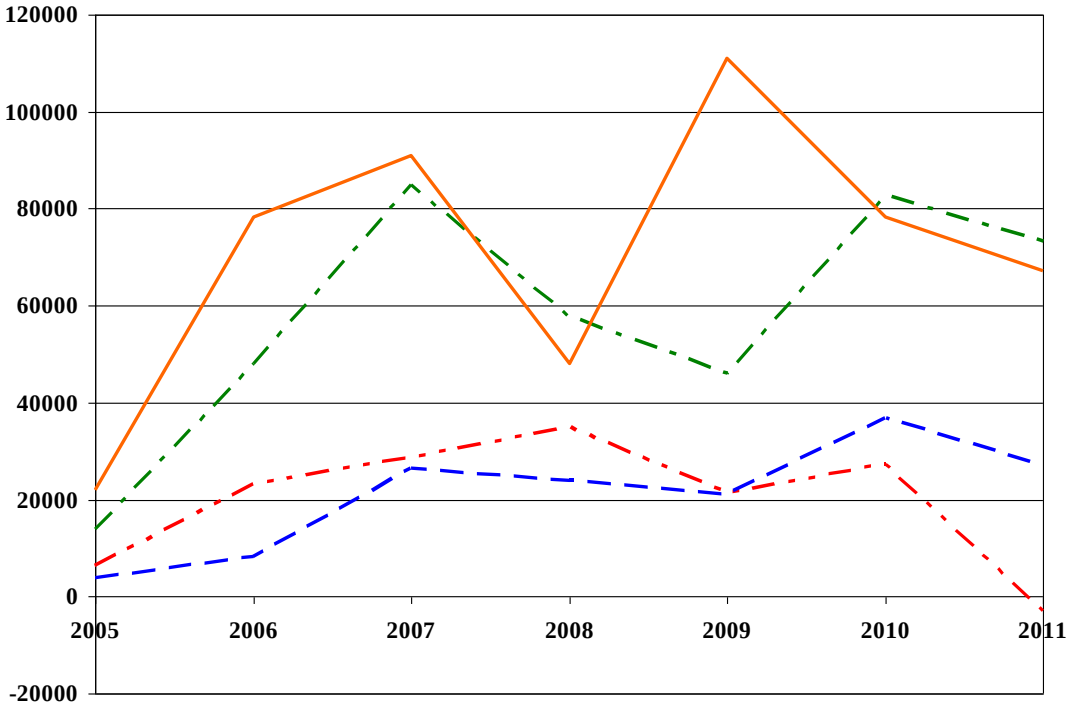


Fig. 2: Estimated Returns by Rainfall Realization

Four ICRISAT Villages 2005-2011

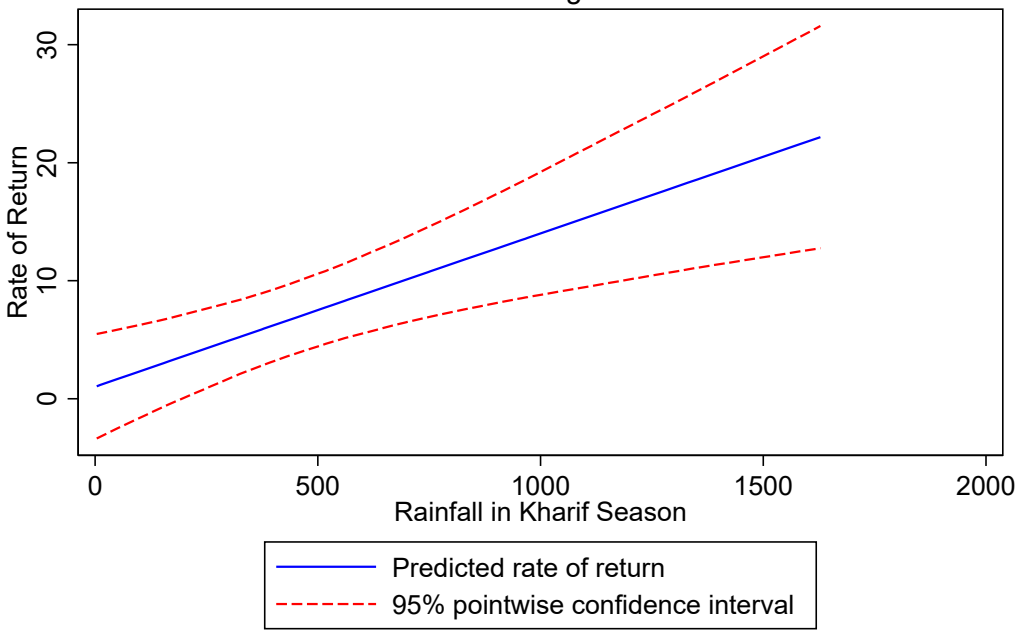
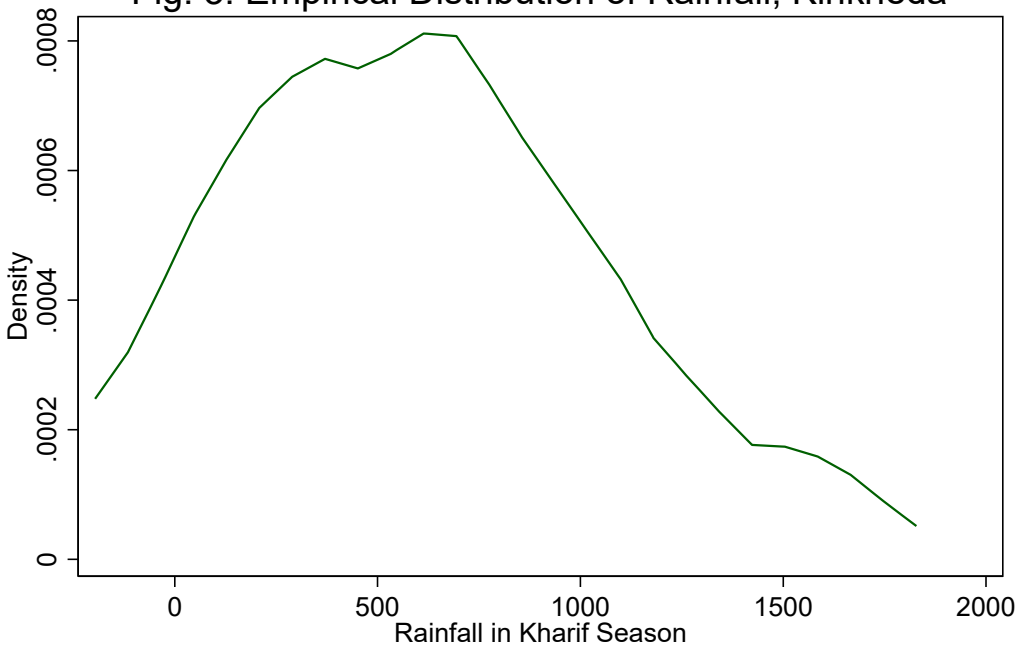
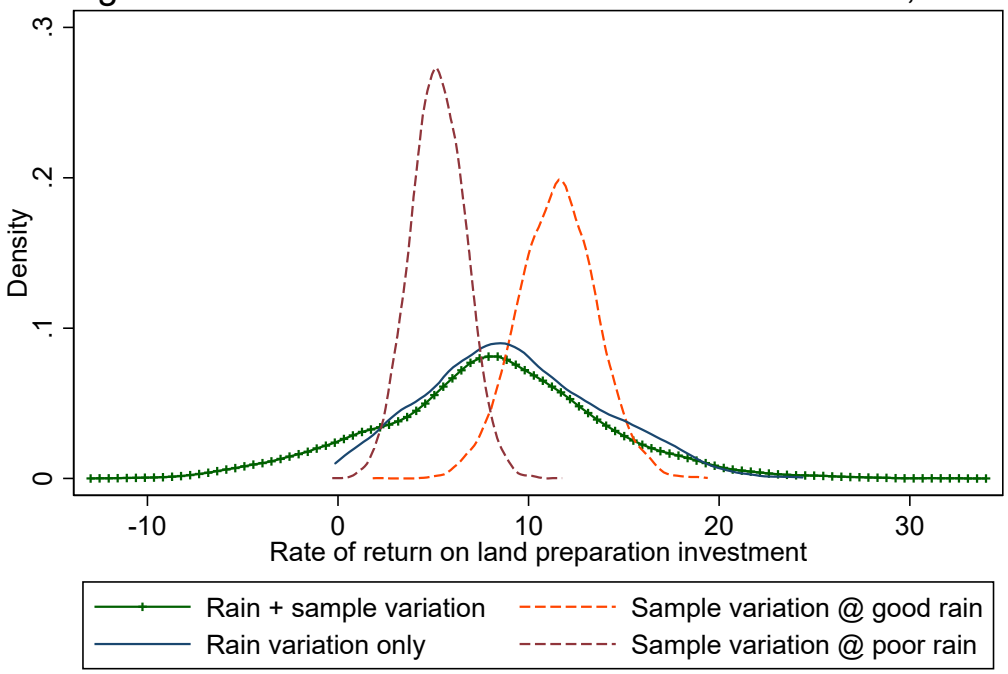


Fig. 3: Empirical Distribution of Rainfall, Kinkheda



Epanechnikov kernel, bw=200

Fig. 4: Distribution of Returns and Weather Shocks, India



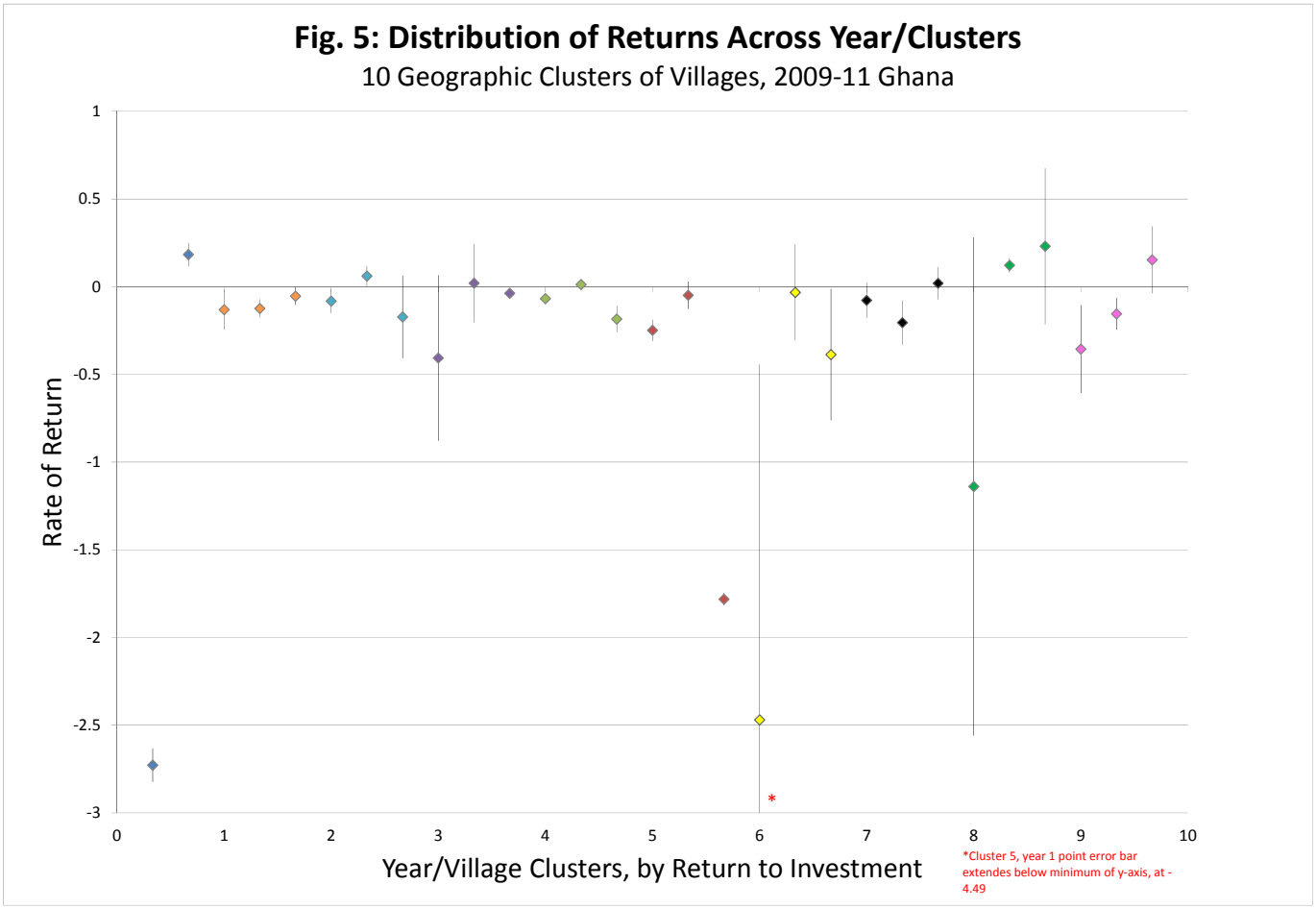
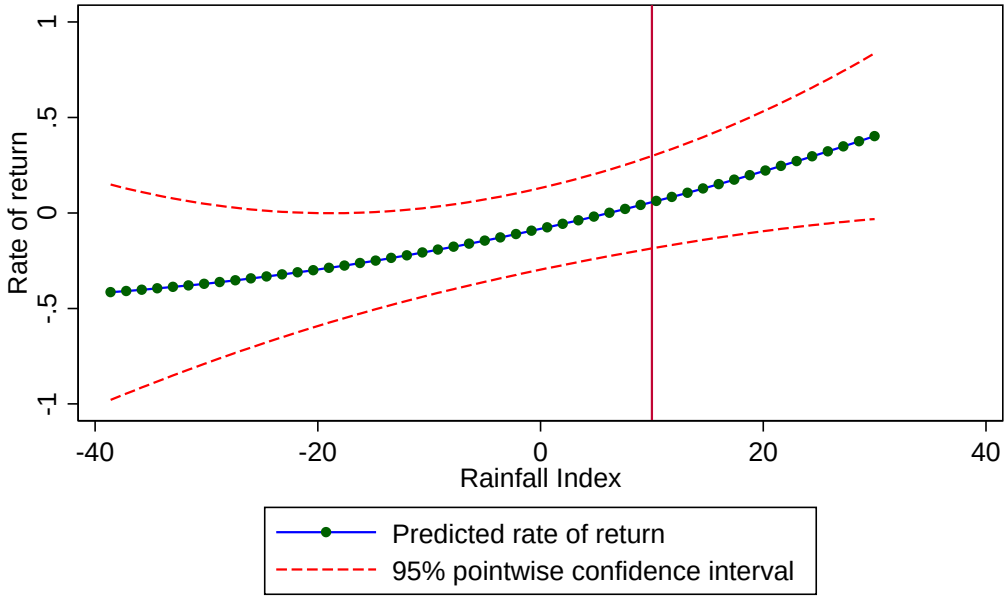


Fig. 5: Distribution of Returns Across Year/Clusters
10 Geographic Clusters of Villages, 2009-11 Ghana
Rate of Return
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
0
1
2
3
4
5
6
7
8
9
10
*
Year/Village Clusters, by Return to Investment
*Cluster 5, year 1 point error bar extendes below minimum of y-axis, at -4.49

Fig. 6: Estimated Returns by Rainfall Realization

75 Northern Ghana Villages 2009-2011



Std. errors clustered at community level. Returns evaluated at median total rainfall.

Fig. 7: Distribution of Returns and Weather Shocks, Ghana

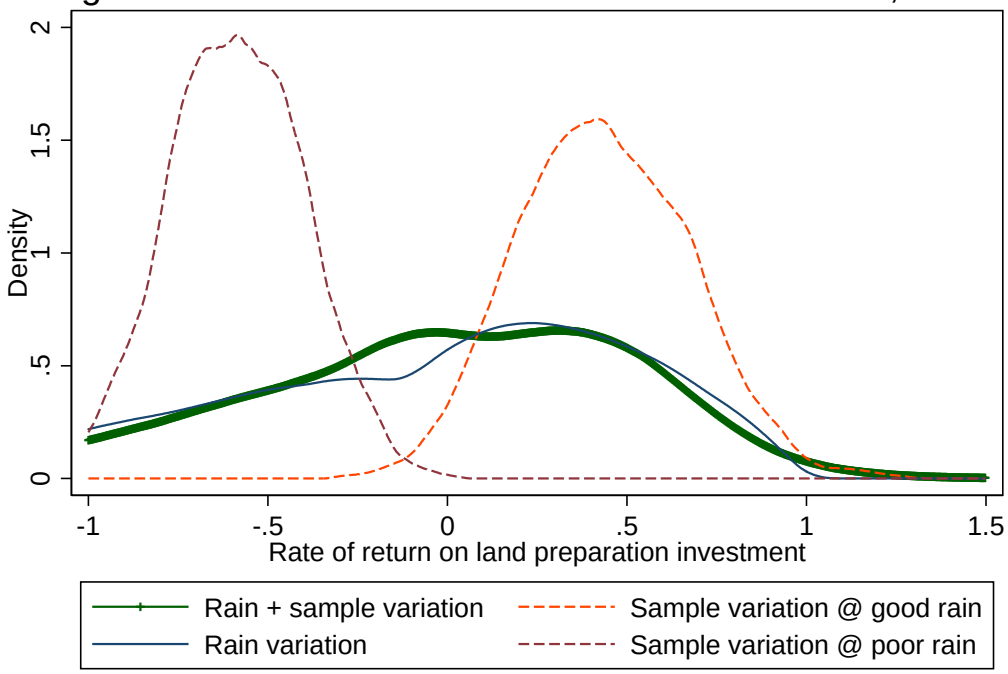
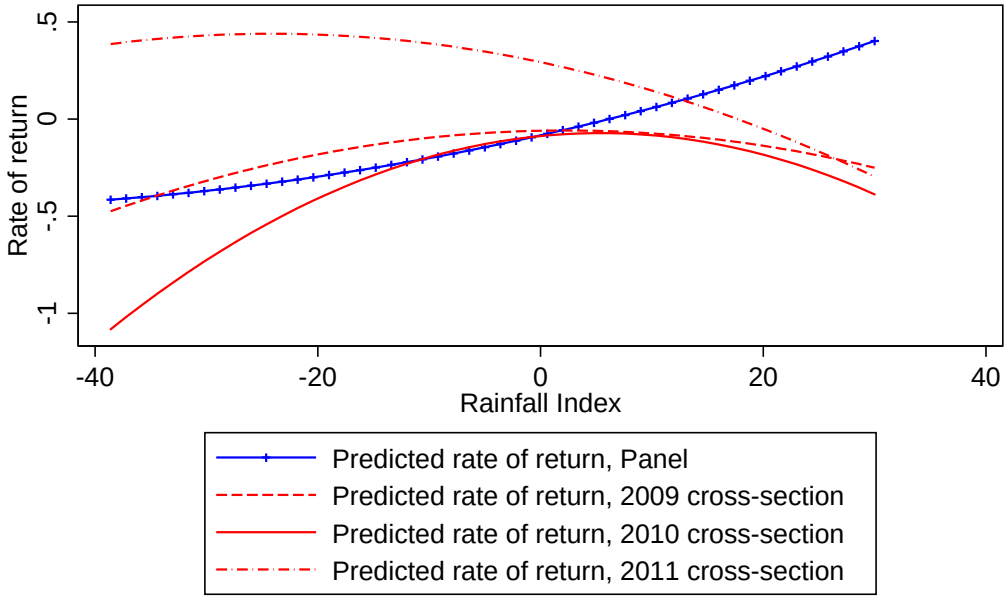


Fig. 8:Estimated Returns by Rainfall Realization

75 Northern Ghana Villages 2009-2011



Std. errors clustered at community level. Returns evaluated at median total rainfall.

Fig. 9: Time Series of Oil Prices and Indonesian Exchange Rate

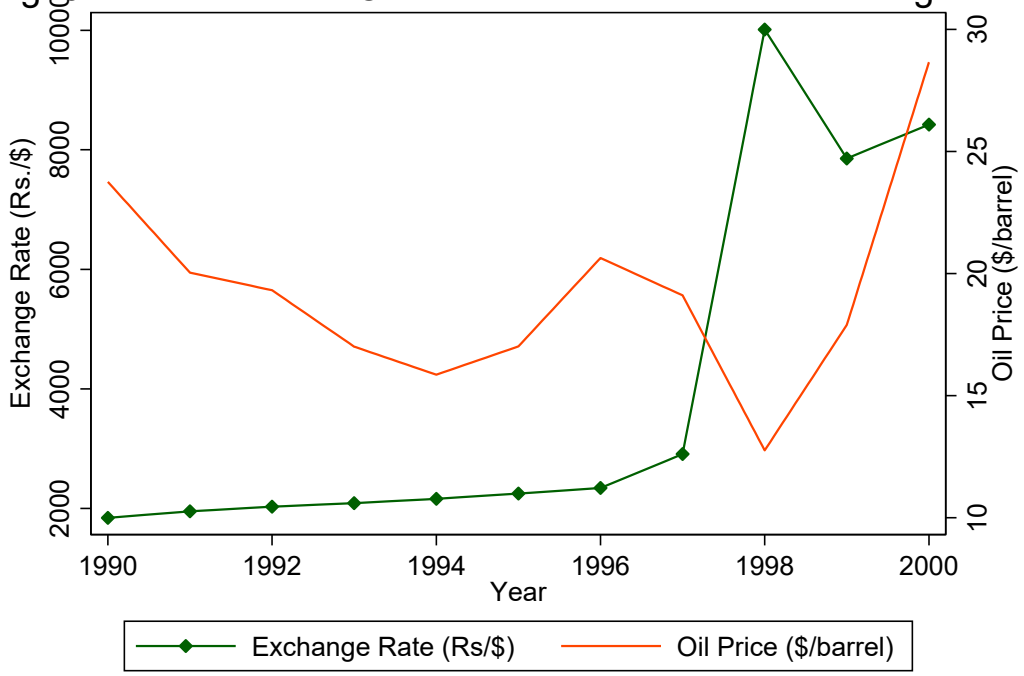
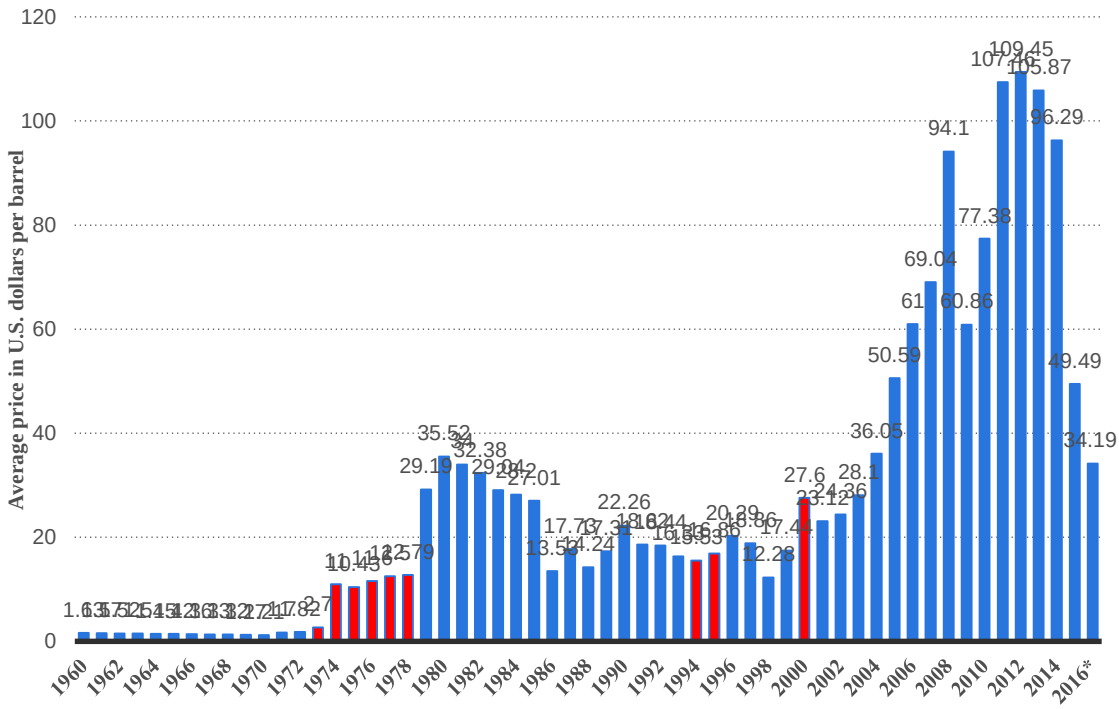


Figure 10. Average Annual OPEC crude oil price: 1960 - 2016
(U.S. dollars per barrel)



Source: OPEC; IEA, Statista, Inc.; ID 262858

Table 1
Farmer Fixed Effects IV Estimates of Planting-Stage Investments on Net Returns,
Four ICRISAT Villages, 2005-2011

Variable	(1)	(2)
Planting-stage investment	8.45 (1.84)	1.53 (2.86)
Planting-stage investment squared x10^{-5}	-8.56 (2.50)	-2.49 (3.48)
Planting-stage investment x *Kharif* rainfall	-	.0163 (.00484)
Planting-stage investment squared x *Kharif* rainfall x10^{-5}	-	-.0160 (.00580)
H_0: Investment effect = 0, ÷^{2}(n) [p]	25.9(2) [.0000]	33.3(4) [.0000]
Investment effect at sample mean investment and rainfall	6.74 (1.39)	7.58 (1.58)
N (farmer-years)	1,152	1,152

Standard errors in parentheses.
Specification includes rainfall and lagged rainfall and rainfall and lagged rainfall interacted with total landholdings, irrigated landholdings, and four soil characteristics; and village*year fixed effects.
The identifying instruments are the year-specific IMD pre-planting monsoon forecast interacted with total landholdings, irrigated landholdings, and four soil characteristics.

Table 2
Probability of Being within ± δ of the Expected Return on Investment,
by δ and Sample

δ (Percentage Points)	ICRISAT	Ghana
10	.017	.09
20	.03	.22
30	.05	.28
40	.07	.34
50	.09	.66

The probabilities are calculated based on equations (3) and (4) in the text. In column 1 we use the estimate of α (the interaction between investment and the rainfall realization, evaluated at mean investment) from Table 1, column 2, multiplied by the estimated standard deviation of the rainfall distribution in the ICRISAT village with the longest rainfall series. In column 2, we use the estimates of α from column 2 of Table 3, evaluated at the mean investment level, and the empirical distribution of the bivariate weather outcomes in the historical rainfall distribution of the Ghana study area.

Table 3
Farmer Fixed Effects IV Estimates of Planting-Stage Investments on Net Returns,
75 Ghana Villages, 2009-2011

Variable	(1)	(2)
Planting-stage investment	.303 (.156)	.948 (.108)
Planting-stage investment squared x10^{-4}	.525 (.113)	.0622 (.0835)
Planting-stage investment x area cultivated	-.00301 (.00337)	-.00316 (.00244)
Planting-stage investment x rainfall index	-	.0134 (.00471)
Planting-stage investment squared x rainfall index x10^{-4}	-	-.00684 (.00408)
Planting-stage investment x rainfall index squared	-	.000113 (.000123)
Planting-stage investment x area cultivated x rainfall index	-	-.0000305 (.0000764)
Planting-stage investment x total rainfall	-	.00137 (.000431)
Planting-stage investment x rainfall index x total rainfall	-	-.0000537 (.0000301)
H_0: Investment effect = 0, ÷^{2}(n) [p]	257.2(3) [.000]	520.2(9) [.000]
Investment effect at sample mean investment and rainfall	.324 (.154)	.606 (.143)
N (farmer-plot-years)	7,546	7,323

Standard errors clustered at the farmer level in parentheses.
The specifications include the rainfall index and its square, five soil characteristics, five slope characteristics, these characteristics interacted with the rainfall index, the distance of the plot to the house, the number of trees, total rainfall, total rainfall and the rainfall index interacted., plot area, and year.
The identifying instruments are randomized offers of index insurance at nine price levels, cash grants, and cash grants plus insurance.

Table 4
Enterprise Fixed Effects IV Estimates of Investments on Net Returns, by Survey Wave
Sri Lankan Micro-enterprises, 2004-2006

Variable	(1)	(2)
Investment (non land)	.0756 (.0242)	-
Entrepreneur sick	-1042.8 (329.4)	-1033.1 (391.9)
Wave 2 investment	-	.133 (.0718)
Wave 3 investment	-	.132 (.0503)
Wave 4 investment	-	.0615 (.0315)
Wave 5 investment	-	.0795 (.0311)
Wave 6 investment	-	.0664 (.0252)
Wave 7 investment	-	.0794 (.0267)
Wave 8 investment	-	.0701 (.0289)
Wave 9 investment	-	.0884 (.0267)
Wave fixed effects	Y	Y
H_0: wave fixed effects=0, $\div^2(8)$ [p]	30.1 [.0002]	5.45 [.709]
H_0: Seasonal pattern, $\div^2(4)$ [p]	-	2.22 [.694]
N (enterprise-quarters)	3,101	3,101

Standard errors clustered at the enterprise level in parentheses.
Sample excludes enterprises directly affected by the 2004 tsunami.
The identifying instruments in column 1 are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male. In column 2 the instruments also include the two instruments interacted by wave indicators.

Table 5
Enterprise Fixed Effects IV Estimates of Investments on Net Returns, by Quarterly Season
Sri Lankan Micro-enterprises, 2004-2006

Variable	(1)	(2)
Investment (non land)	-	.0977 (.0216)
Entrepreneur sick	-1066.5 (366.5)	-951.5 (345.3)
Q1 investment	.0947 (.0219)	-
Q2 investment	.1001 (.0236)	-
Q3 investment	.0693 (.0192)	-
Q4 investment	.0836 (.0242)	-
Season fixed effects	Y	Y
H_0: Season fixed effects = 0 $\div^2(4)$ [p]	9.27 [.0547]	11.6 [.0208]
N (enterprise-quarters)	3,101	3,101

Standard errors clustered at the enterprise level in parentheses.
Sample excludes enterprises directly affected by the 2004 tsunami.
The identifying instruments in the first column are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male and those two instruments interacted with wave indicators. The identifying instruments in column 2 are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male and quarter indicators.

Table 6
Log Urban Hourly Wages: Estimates of the Returns to Schooling by Year, 1993-2000
Men Aged 16-79 in Urban Indonesia (IFLS 1993, 1997, 2000)

Variable/Estimation Procedure	OLS	OLS	FE
Years of schooling	.123 (.00485)	-	-
Years of schooling/ Difference from 2000, 1993	-	.154 (.0103)	-.00360 (.0160)
Years of schooling/ Difference from 2000, 1994	-	.124 (.00730)	-.0129 (.1013)
Years of schooling/ Difference from 2000, 1995	-	.121 (.00692)	-.0139 (.0102)
Years of schooling/ Difference from 2000, 1996	-	.122 (.00677)	-.0125 (.0102)
Years of schooling/ Difference from 2000, 1997	-	.138 (.00768)	-.000132 (.0104)
Years of schooling/ Difference from 2000, 1998	-	.104 (.00651)	-.0255 (.00752)
Years of schooling/ Difference from 2000, 1999	-	.102 (.00623)	-.0270 (.00707)
Years of schooling 2000	-	.129 (.00761)	-
Year fixed effects	Y	Y	Y
H_0: Year fixed effects=0, F(8, 3062) [p]	415.0 [.0000]	395.1 [.0000]	17.0 [.0000]
H_0: Schooling returns are equal across years, F(7,3062) [p]	-	5.94 [.0000]	2.95 [.0044]
N (person-years)	12,769	12,769	12,769

Standard errors clustered at the individual level in parentheses.
Specifications also include age and age squared.

Table 7
Log Urban Hourly Wages: Estimates of the Returns to Schooling and External Macro Shocks
Men Aged 16-79 in Urban Indonesia (IFLS 1993, 1997, 2000)

Sample Period	1993-2000		1998-2000	
Estimation Procedure	OLS	FE	OLS	FE
Years of schooling	.0855 (.0102)	-	.00134 (.0213)	-
Years of schooling x oil price	.00105 (.000492)	.00156 (.000434)	.00214 (.000517)	.00210 (.000631)
Years of schooling x exchange rate x industry openness x ind. skill x 10^{-5}	-	-	.222 (.0506)	.121 (.0615)
Oil price	-.0528 (.0136)	-.0504 (.0106)	-.0894 (.0131)	-.0632 (.0196)
Exchange rate x industry openness x skill x10^{-4}	-	-	-.342 (.107)	-.613 (.258)
Exchange rate x10^{-4}	-	-	.447 (.209)	.437 (.746)
Industry openness x skill	-	-	.0968 (.0533)	-
H_0: No shock effects on schooling returns, F(2, 2120) [p]	-	-	13.3 [.0001]	5.54 [.0040]
N (person-years)	12,242	12,242	5,470	5,470

Standard errors clustered at the individual level in parentheses. Specifications also include age and age squared, year, and age and age squared interacted with the oil price/industry variables.