The agricultural wage gap within rural villages

Ceren Baysan, Manzoor H. Dar, Kyle Emerick,
and Elisabeth Sadoulet*

September 4, 2019

Abstract

We use unique data on daily labor-market outcomes for Indian casual workers to study labor reallocation between agricultural and non-agricultural activities *within rural areas*. We find that workers who switch sectors during a period of one to two weeks can obtain 21 percent higher wages by taking nearby non-agricultural jobs. Our estimate is unlikely to be explained by common mechanisms from the literature, such as sorting on worker ability, migration costs, or disamenities associated with living in urban areas. Put differently, new explanations are needed for the sectoral wage gaps that exist within rural areas.

*Baysan: University of Essex, ceren.baysan@essex.ac.uk; Dar: seedmanzoor.dar@gmail.com; Emerick: Tufts University and CEPR, 8 Upper Campus Road, Medford, MA 02155-6722, kyle.emerick@tufts.edu; Sadoulet: University of California at Berkeley, e-mail:esadoulet@berkeley.edu. We acknowledge financial support from the J-PAL & CEGA's Agricultural Technology Adoption Initiative (ATAI) and from the Stress-Tolerant Rice for Africa and South Asia (STRASA) project of the CGIAR. Emerick is grateful to the Institute of Economic Development at Boston University where he was a visiting scholar while part of this research was carried out. Baysan is grateful to the Weatherhead Center for International Affairs at Harvard University where she was a postdoctoral fellow while part of this research was carried out and gratefully acknowledges financial support from the NSF Graduate Research Fellowship program.

1 Introduction

Most people in poor countries work in agriculture. As examples, agriculture accounts for nearly half the labor force in India and more than 70 percent across several countries in Sub-Saharan Africa.[1] At the same time, agricultural wages fall below wages in other sectors (Restuccia, Yang, and Zhu, 2008; Gollin, Lagakos, and Waugh, 2014). On the one hand, this could imply a misallocation of workers across sectors and therefore an opportunity for policy that transitions workers out of agriculture — usually by encouraging rural-urban migration. Sorting of workers by ability provides an alternative explanation. Perhaps high-ability workers leave the farm for cities while low-ability workers stay in agriculture (Young, 2013; Lagakos and Waugh, 2013). Recent empirical evidence from multiple countries aligns with this latter explanation: agricultural wage gaps appear much smaller — or disappear entirely — when controlling for different measures of individual ability (Hamory Hicks et al., 2017; Alvarez, 2018; Herrendorf and Schoellman, 2018).[2]

We estimate the wage gains from transitioning out of agriculture using a unique dataset on labor market outcomes for a panel of Indian agricultural workers. Our dataset has two elements that allow us to build on the previous work in this area. First, we observe daily wages for workers changing sectors within a short time window of just one to two weeks. This allows us to control for time-varying attributes of individual ability that may vary over a longer time period, but are plausibly unchanged during a shorter period, such as vocational training. Second, unlike the previous work on rural-urban migrants, the non-agricultural work we observe also takes place in rural areas, often in the respondent's own village. Separating non-agricultural work from rural-urban migration allows us to rule out many of the possible explanations from the literature for an agricultural wage gap. For instance, geographic differences between locations cannot explain differences between agricultural and non-agricultural wages in our setting.

Moreover, understanding the wage gap between the two sectors within rural areas is important in its own right and has received little attention. 69% of the Indian population was living in rural areas in 2011. At the same time, the rural non-agricultural sector has become an increasingly significant source of employment. For example, the share of rural male workers employed in non-agricultural work increased from 22% in 1983 to 41% in 2011 (Binswanger-Mkhize, 2013; D Narasimha Reddy and Bantilan, 2014).

[1]Data from the FAO — which are usually based on either household surveys, labor force surveys, or population censuses — show that the agricultural labor share in India is 49.7 percent. Uganda (71.9), Ethiopia (72.7), Guinea (74.8), Rwanda (75.3), and Burkina Faso (78.4) are examples of Sub-Saharan African countries with agricultural labor shares over 70 percent.

[2]Sorting on ability also explains spatial wage differences in other contexts. Combes, Duranton, and Gobillon (2008) find that it explains a large share of the spatial variation in wages for French workers.

We find that workers earn higher wages by moving from agricultural to non-agricultural work. Sorting on individual attributes is an unlikely explanation for this wage gap. We arrive at this finding using detailed data on labor allocation for a panel of agricultural workers during the 2014-2016 cropping seasons.[3] About 21 percent of workers switch sectors during the periods when we observe wages in both sectors. Using these workers for identification, by including individual and survey fixed effects into the regression, we estimate a within-individual agricultural wage gap of 21 percent. Put differently, the same worker can obtain 21 percent higher wages by moving out of agriculture and into non-agricultural work.

Some components of unobserved ability may vary over time. A worker may gain additional skills in between agricultural seasons. Or they may migrate in response to a time-varying change in human capital. These changes may occur over several months, but are unlikely to happen over a period of just a few days. Using our daily observations on wages, we find the same wage gap when including individual-by-survey round fixed effects. In other words, switching sectors within a one to two week period generates the same wage gains as switching over a longer time horizon. This estimation confines the identification to fewer individuals, but helps rule out all of the unobserved worker attributes that remain constant within such a short period.

This result differs from the recent literature using rural-urban migrants and finding that sorting on unobservable ability explains the gap between rural (agricultural) and urban (non-agricultural) wages. Our data point in a different direction for the gap between agricultural and non-agricultural wages within rural areas: a meaningful gap persists even after eliminating the sorting explanation.

We then explore other possible reasons for this agricultural wage gap. The nature of our data allows us to rule out common explanations for rural-urban wage gaps from the literature. The wage gap does not reflect migration costs. Rather, workers can obtain higher wages by taking nearby non-agricultural jobs, i.e. those in the same village or nearby villages that don't require fixed migration investments. Consistent with this argument, we find similar results when dropping non-farm work outside of the respondent's own village. Another common explanation is that people may forgo higher non-agricultural wages because they prefer amenities in rural over urban areas (Imbert and Papp, 2018; Lagakos, Mobarak, and Waugh, 2018). Confining our estimation to rural areas, and focusing on daily transitions between sectors makes location-specific explanations like this one seem unlikely.

Perhaps search costs make even nearby non-agricultural work difficult to find, or some

[3] The data were collected as part of a randomized evaluation of the effects of a new drought-tolerant rice variety on labor markets. The technology was introduced in 2014 and we collected the six followup phone surveys during the planting and harvesting times for that season and the following two seasons.

people are unable to perform the tasks required for non-agricultural work. We consider these possibilities by testing whether workers move into the non-agricultural sector when they have fewer options because agriculture faces a bad year, i.e. their own farms are less productive, and agricultural employment is more scarce. Using variation in monsoon rainfall as a measure of agricultural productivity, we find that low agricultural productivity leads to a sharp increase in rural non-agricultural work. Going from the 90th to the 10th percentile of the rainfall distribution, conditional on village and year fixed effects, causes rice yield to decline by 63 percent, the probability of working in agriculture at harvesting to decline by 8.5 percentage points (39 percent), and the probability of working in the non-agricultural sector to increase by 6.2 percentage points (39 percent).[4] In short, people take rural non-agricultural jobs when agricultural opportunities are less available. This finding suggests that the inability to find non-agricultural jobs, or the lack of ability to perform these tasks are unlikely explanations for the agricultural wage gap.

Taken together, these findings suggest that the common explanations for agricultural wage gaps cannot fully explain our results. To gain additional insight into other possible mechanisms, we posed a simple question to workers: what is the top reason you work in agricultural jobs if wages in those jobs are a bit lower than non-agricultural jobs? Workers give multiple explanations. But, the most frequent explanation is that non-agricultural jobs are "too hard." We interpret this finding as suggestive that the agricultural wage gap partly reflects compensation for the difficulty of rural non-agricultural work — which in this setting is indeed physically demanding. The available jobs tend to involve construction, brick laying, and working in brick factories or coal mines. Workers prefer not to take these jobs, but do so when there are fewer options in agriculture. One limitation of this exercise is that we learn from self-reported explanations from workers, rather than their revealed behavior. Nonetheless, this plausible mechanism highlighted by workers — that agricultural wage gaps reflect compensating differentials for the type of work being done — deserves consideration as a possible explanation for why workers earn more outside of agriculture.[5]

Our paper adds to the literature on labor reallocation and development. This literature — focusing on rural-urban migration — seeks to explain the large gap between rural and urban wages in developing countries. The list of explanations includes sorting on ability (Hamory Hicks et al., 2017; Alvarez, 2018), the costs of moving across space (Bryan and

[4] Colmer (2018) uses data across all of India and similarly shows that people turn to non-agricultural work when temperature is unfavorable for agriculture. His district-level estimates include movement into the non-agricultural sector while continuing to reside in the village, and short-term migration to district towns.

[5] Compensating wage differentials have been investigated in several other contexts in labor economics (Smith, 1979). As one example, Duncan and Holmlund (1983) look at how changes in job attributes relate to changes in wages amongst adults in Sweden.

Morten, 2017), and the disamenities that come with living in urban areas (Munshi and Rosenzweig, 2016; Bryan and Morten, 2017; Lagakos, Mobarak, and Waugh, 2018; Morten, 2019).[6] Our paper is the first to use high-frequency data in order to estimate a sectoral wage gap within rural areas — a setting where the above mechanisms explaining rural-urban wage gaps do not operate. Therefore, our finding indicates that additional explanations are needed to explain why rural non-agricultural wages remain higher in equilibrium.

The remainder of the paper is organized as follows. Section 2 describes our data and how it allows us to control for additional unobserved worker attributes while at the same time ruling out some of the common explanations for the rural-urban wage gap. This section also outlines our estimation approach. Section 3 discusses our results and Section 4 offers brief conclusions.

2 Data and Methods

2.1 Data and descriptive statistics

Our primary sample is spread across 12 blocks within 4 districts of the Jharkhand state in eastern India. The blocks were identified as being suitable for a drought-tolerant rice seed variety that we were testing using a randomized controlled trial. We selected a random sample of villages amongst those with 30 to 550 households. Within each village, enumerators located a village leader and asked for names of 35 people from separate households: the 25 largest rice farmers, 7 male individuals that work on other farmer's fields, and 3 female individuals that also work as casual agricultural laborers. Enumerators carried out a baseline survey with the farmers and workers during the period from late April to early June 2014.

Our sample of laborers consists of people that are landless or have small amounts of land. This population makes up a non-trivial share of the people dependent on agriculture in rural India. In contrast to large landowners, these workers generate most of their income from supplying labor to the casual labor market.

Hiring and wages in casual labor markets in India are generally determined on a daily basis. Yet, most studies rely on data that aggregates labor market outcomes over a longer time period. This potentially misses short-term movement between occupations. To better measure labor-market outcomes in our context, we collected daily data on wages and employment. We did this by conducting phone surveys that took place during the transplanting and harvesting periods across the 2014, 2015, and 2016 cultivation seasons. Wet-season rice

[6] Focusing on Sub-Saharan Africa, Gollin, Kirchberger, and Lagakos (2017) find less evidence for a negative correlation between the quality of amenities and population density.

is the dominant crop in our sample area and is planted in late July / early August and is harvested in late November. Our phone surveys took place during these times to coincide with the times of the year when agricultural labor is employed. There is very little agricultural work during other times because lack of irrigation limits dry-season cultivated area.

During the first year (August and November 2014) surveyors attempted to contact the 10 laborers in each of the 200 villages. During each call respondents were asked whether they worked on another person's farm or their own farm, the wage they received, whether the work took place in their own village, and their activity if they did not work in agriculture. This information was collected for the seven days preceding the phone call. We repeated this same process in the 2015 and 2016 seasons with a few important differences. First, we expanded the sample to include 6 female laborers per village. The additional three laborers were selected from a census that had been conducted in all villages on households with casual laborers.[7] Second, starting with the 2015 harvesting survey, we expanded the recall window to 14 days to more easily capture the entire planting or harvesting period for each village. The phone surveys produced a high response rate: an average of 86 percent of the workers in the baseline sample were reached.[8]

These data allow us to observe daily employment outcomes for planting and harvesting across three agricultural seasons. In addition, we collected non-agricultural wages in the 2015 planting and both 2016 surveys. These observations consist mostly of casual work for a daily wage — rather than self employment. We observe the daily wage for 82 percent of the non-agricultural work days in these three surveys. This information, along with the individual-level panel on agricultural outcomes, allows us to measure the agricultural wage gap while controlling for unobserved heterogeneity across individuals.[9] Since the people switching sectors give identification, it is useful to compare them to the individuals that work in agriculture for the entire sample period. About 20 percent of the workers from the baseline survey switched sectors. Table 1 shows the differences between these two groups. Switchers are predominantly male and generally poorer in several dimensions. For example, they are less likely to have access to electricity, more likely to be in households using the

[7]We discovered after looking at our first year of data that our sample of laborers was under-representative of females based on their importance as agricultural workers. In addition to adding more females to the sample, we make use of data on hiring from farmers to weight our worker data by gender. We do this to make our labor-market outcomes representative of an average agricultural worker. Section 2.2 provides details on the gender weights.

[8]The response rate ranged from 79 percent in the third year planting survey to 91 percent in the year two planting survey.

[9]Our main specification includes individual and survey round fixed effects. These estimates could be affected by unobserved time-varying attributes, such as changes in ability or training. We also include a specification with individual-by-survey round fixed effects. Unlike longer term changes, switching sectors within one to two weeks is less likely to be correlated with a time-varying change in ability or training.

government's rural employment guarantee (NREGS), have larger households, and more likely to belong to lower castes. They are also more likely to have household members that migrate temporarily (outside the village), but are not more likely to engage in permanent migration. Yet, switchers have no less land. The average laborer household cultivates 0.57 acres during the rainy season and only about 16 percent of households cultivate no land at all.[10] Overall, the people that switch between local agricultural and non-agricultural work are neither the wealthiest or most educated. If anything, the switchers tend to come from poorer households.

Figure 1 further describes our data by showing a breakdown of the activities we observe. About 30 percent of the sample work only on their own farms during planting and harvesting. About 25 percent split time between agricultural wage labor and own-farm work, while another 25 percent of workers only engage in agricultural wage labor. The workers that switch sectors during the same survey round constitute about 4 to 8 percent of each survey. We use these workers to verify that our main estimate changes little when confining the identification to this smaller group.

We also make use of three additional sources of data. First, we surveyed the 10 largest farmers after harvesting in each of the three years.[11] We use these farm-level data to characterize the variation in agricultural productivity across our sample and to understand how shocks to agricultural labor demand affect non-agricultural employment. Second, to measure these shocks, we use daily rainfall estimates from the Climate Hazards Group InfraRed Precipitation with Station dataset (CHIRPS) (Funk et al., 2015). CHIRPS incorporates 0.05° resolution satellite imagery with station-level data to create a gridded daily time series. We calculated average precipitation across the 200 sample villages to generate a daily average precipitation for the sample area. Figure A1 helps visualize these data. It shows that 2014 and 2015 — the first two years of our data collection — were drier years with particularly long dry spells during the growing season. In contrast, 2016 was the wettest year since 2000. The productivity data from farmers highlight the importance of timely rainfall. Relative to 2016, yields were lower by 56 percent in 2015 and 25 percent in 2014. Third, since the numbers of male and female workers to include in the panel surveys were arbitrarily chosen, we use phone surveys with farmers on labor demand to calculate the correct gender-specific weights for our sample of laborers.

[10]The average cultivated area of the laborer households amounts to about 20 percent of the average cultivated area of the sample of large farmers.

[11]These farmers were selected amongst the 25 farmers listed at the beginning of the study.

2.2 Empirical Approach

We observe $wage_{ivtd}$, which is the wage for worker i, residing in village v, during survey round t and on day d. The daily data permit us to estimate the wage gap between agricultural and non-agricultural work. To do so, we estimate,

$$log(wage_{ivtd}) = \alpha_{iv} + \delta_t + \beta NonAg_{ivtd} + \varepsilon_{ivtd}, \tag{1}$$

where $NonAg_{ivtd}$ is an indicator for wage labor in the non-agricultural sector, α_{iv} is an individual fixed effect, δ_t is a survey round fixed effect, and ε_{ivtd} is an error term that we cluster at the village level. We limit the data for this estimation to the three survey rounds where we collected wages in both sectors. The parameter β measures the wage differential on days when a given worker took a non-agricultural job with those when agricultural wage labor was selected. The individual fixed effect eliminates time-invariant individual attributes. We also check a stricter specification with individual-by-survey round fixed effects. Previous work on rural-urban migration has estimated sectoral wage gaps using people who switch sectors over longer time periods (Hamory Hicks et al., 2017; Alvarez, 2018; Herrendorf and Schoellman, 2018). Our specification with the shorter time window allows us to estimate a gap within rural areas for jobs that can be taken within a period of just one to two weeks.

The phone surveys with farmers indicate that about 82 percent of the workers hired during years 2 and 3 are females, which is larger than the proportion of females we selected in our sample.[12] We therefore weight the observations in the analysis. For each of the survey rounds, the weight for female observations is calculated as the share of the hired workers that are female — across all of our phone surveys with farmers — divided by the share of respondents from that survey wave that were female. We define the weights analogously for males. Although not affecting our results, this weighting scheme ensures that our estimates represent the average casual agricultural worker.

3 Results

3.1 The agricultural wage gap

Table 2 shows our estimates of the agricultural wage gap. Our main estimate in column 1 is identified off of about 21 percent of the sample and includes individual and survey fixed effects. It shows that the same individual increases his/her daily wage by 21 percent when

[12] Part of the reason for this is that our phone surveys collected information during planting and harvesting — two activities more likely to be done by females. Males are more active during land preparation (plowing) and post-harvest activities like crop threshing.

moving out of agriculture and into non-agricultural work.[13] This estimate, however, is partly identified off of people switching sectors across survey rounds. Time-varying unobservable attributes, such as changes in skills or physical health, could drive part of the estimate. Reassuringly, column 2 shows that we obtain the same result with individual-by-survey round fixed effects. This confines the identification to about half as many individuals, but we estimate the same wage gap of 21 percent. The types of unobservables that might vary over the period of a few months, but are less likely to vary over the period of one to two weeks, do not appear to drive our estimate.

Columns 3-5 show the unadjusted agricultural wage gaps where we do not include individual fixed effects. Non-agricultural wages are higher by about 30 percent compared to agricultural wages — regardless of whether we use variation within or across villages. In our case, individual attributes explain only about a third of the wage gap. Unlike the literature on rural-urban migrants, which finds that this type of selection accounts for most of the rural-urban wage gap, we find that much of the wage gap remains when conditioning on individual fixed effects. As an additional note, none of the estimates in Table 2 change meaningfully if we do not to weight the observations by gender (Table A1).

To put our estimate in context, Herrendorf and Schoellman (2018) use census data across 13 countries to show that non-agricultural wages are about 1.8 times higher than agricultural wages. Adjusting for only human capital, gender, and geographic location causes their estimate to decrease to 1.33. Our estimates suggests that focusing on the rural non-agricultural sector, and eliminating the most plausible unobserved correlates of ability, the non-agricultural gap remains at about 1.23.[14]

3.2 Possible explanations

Why does the agricultural wage gap exist? Besides sorting on unobservables, several mechanisms have been pointed to in the literature. These include differential amenities, limited mobility, search costs, or limited availability of non-agricultural work. We next investigate the likelihood of each of these possible mechanisms in our data.

> **Lost amenities when taking non-agricultural work:** Leaving the farm may cause one to lose valued amenities such as quality housing (Lagakos, Mobarak, and Waugh, 2018) or access to traditional risk-sharing arrangements (Munshi and Rosenzweig, 2016). These disamenities are most often associated with rural-urban migration.

[13] In line with the descriptive evidence above, only about 15 percent of these non-agricultural work days are from females.

[14] The precise gap from the log wage regression is $e^{0.207} = 1.23$.

For instance, Imbert and Papp (2018) find that rural-urban migration in India can lead to higher wages, but the non-monetary costs of being in the city are enough to dissuade potential migrants. In contrast, the non-agricultural work we analyze is within rural areas and can be completed within the same 7-14 days as working in agriculture, making it seem unlikely that these lost amenities can account for the gap between wages in the two sectors. Thirty-one percent of the non-agricultural wage observations we use in Table 2 were from activities outside of the worker's village. We find that dropping these observations and re-estimating the wag gap leads to similar results (Table A2).

Costs of spatial mobility: Distance from home may be a reason why workers avoid non-agricultural work. This could be because more non-agricultural work becomes available when moving further outside of rural villages, or because non-agricultural work is associated with travel to towns or cities. Despite this, mobility does not seem to be the limiting factor behind the agricultural wage gap we estimate. As further evidence, 54 percent of workers (124 out of 230) that switch sectors do so on back-to-back days. We therefore observe 162 days in which a worker transitioned between agricultural and non-agricultural work on successive days. It is unlikely that this type of switching between jobs would require large transport costs or lead to any loss of rural amenities. We nevertheless find a similar wage gap in these daily transitions: Relative to days without changing of sectors, we find that moving from agricultural to non-agricultural work increases wages by 17.6 percent and a transition in the reverse direction decreases wages by 28.6 percent.[15]

Our data on agricultural employment also include whether the work took place in another village. Table A3 shows that there is a wage premium for leaving the village, however it is only about 4 to 5 percent. The magnitude of this effect is small compared to the non-agricultural wage premium — especially since non-agricultural work only sometimes requires leaving the village. These data provide further evidence suggesting that travel costs are not the key factor behind the agricultural wage gap.

Search costs and limited availability: Non-agricultural work may be difficult to find or not available to many workers. These search costs could explain the persistence of the agricultural wage gap. If there was a binding constraint on the availability of non-agricultural work, then obtaining non-agricultural jobs would be even more difficult when employment opportunities in agriculture fall and hence more workers

[15] These estimates come from a regression where the change in log daily wages ($log(wage_{ivtd}) - log(wage_{ivtd-1})$) is regressed on two dummy variables: one for a transition from agriculture to non-agriculture and another for a transition from non-agriculture to agriculture.

seek to find non-agricultural work. We thus test whether workers are less likely to find non-agricultural work following a negative shock to agricultural labor demand. We focus on the three harvesting surveys and estimate

$$employment_{ivtd} = \gamma_t + \alpha_v + \beta Rainfall_{vt} + \varepsilon_{ivtd}, \tag{2}$$

where $employment_{ivtd}$ is one of four indicator variables for working as an agricultural wage laborer, supplying labor to their own farm, working in the non-agricultural sector, or leisure / housework. Harvesting takes place in November or early December. Therefore, we calculate cumulative village-level rainfall from June through October to measure shocks to agricultural labor demand. Showing that workers easily replace agricultural work with non-agricultural jobs would suggest that search costs and availability of non-farm work are not limiting.

Figure 2 presents the results visually by first residualizing the data to eliminate the fixed effects, and then showing binned scatter plots of different outcome variables against rainfall realizations. The precise parameter estimates are in Table 3. As expected, agricultural productivity increases with higher rainfall. The upper-left panel of the figure shows a tight positive association between total precipitation and rice yield: going from the driest to the wettest observations causes yield to more than double. The remaining panels in the figure show how the daily earnings and time allocation of casual laborers *at the time of harvesting* respond to these rainfall shocks. Daily earnings from agriculture and the likelihood of obtaining agricultural employment both decrease with low rainfall. Similarly, laborers are less likely to spend time working on their own fields during lower-rainfall years. On the other hand, non-agricultural employment at harvesting increases during drier years and falls during good agricultural years. Workers are also more likely to report "doing nothing" or being engaged in housework with low rainfall.

These findings suggest that laborers can obtain very local non-agricultural work when they want. Put differently, the type of non-agricultural jobs that we show lead to higher wages are not inaccessible. Laborers move into non-agricultural work when they are less able to find work in agriculture.

Does the rise in non-agricultural work during low-rainfall years result directly from an influx of workers unable to work in agriculture? In contrast to this shift in non-agricultural labor supply, an alternative explanation would be that demand for non-agricultural work is increasing during dry years, for instance if rainfall hinders construction. We use data on non-farm wages to distinguish these two possibilities. While our phone surveys only collected non-agricultural wages during harvesting for the 2016 season, we also have observations on non-agricultural wages during the harvesting period from the followup survey with workers after the 2014 season. Using these two years of data, Figure 3 shows a positive relationship between local rainfall and non-agricultural wages, net of village and year fixed effects. Combining this with the non-agriculture employment panel in Figure 2, poor rainfall (for agriculture) leads to higher non-agricultural employment and *lower* non-agricultural wages. Wages and employment moving in opposite directions is more consistent with workers increasing their supply of non-agricultural work when agricultural employment is less available.

Worker explanations for not taking non-agricultural work: In the final followup survey after the 2016 harvest, we posed a simple question to our sample of laborers: why would you continue to work in agriculture if wages are lower there compared to non-agricultural jobs? While not based on revealed behavior, responses to this question give some insights into what might be behind our estimated wage gap.

Responses vary, but Figure 4 shows that the top answer is that non-agricultural jobs are "too hard". Twenty-three percent of workers point to this as a reason for not wanting to close the wage gap between sectors. This evidence does not pinpoint what exactly makes non-agricultural jobs harder. It instead provides suggestive evidence that workers prefer a day of agricultural work over local non-agricultural employment. This could be because non-agricultural jobs are more physically demanding, require longer hours, or involve tasks that are less familiar than agricultural activities.[16] Indeed, non-agricultural work in rural areas often requires physically demanding tasks. During this same survey we asked workers what they do when working in non-agricultural jobs. These jobs involve some form of construction around 68 percent of the time. Other popular activities include working in local coal mines or brick kilns.

[16]The preference for agricultural work remains puzzling even if non-agricultural employers require longer days. It indicates that workers would prefer to earn less in a day in exchange for continuing to work in agriculture — even when they spend many other days without wage employment, i.e. working on their own very small farms or doing household chores. Our 2014 follow up survey includes information on the length of the agricultural work day. Farmers report an average agricultural work day of 7.7 hours for males and 7.5 hours for females. Using variation in daily hours, Table A4 shows that daily wages are not positively correlated with the length of the working day. These data suggest that the relevant unit for wage determination is the day, rather than the hour.

4 Concluding Remarks

Models of labor (mis)allocation in developing countries tend to focus on reallocation across space from rural to urban areas. Reallocation across sectors within rural areas has received less attention. In this paper we have shown evidence that laborers in rural Indian villages can increase daily earnings by about 21 percent from moving out of agriculture and working in the local non-agricultural sector.

We cannot definitively point to a single explanation for this rural agricultural wage gap. But our data allow us to provide some evidence that it cannot be explained by the common explanations for the rural-urban wage gap. In particular, we cannot explain the agricultural wage gap with sorting on unobservable ability. Similarly, different amenities or other utility costs of migration cannot explain why workers choose not to take nearby non-agricultural jobs that do not require migration.

Direct surveys with workers reveal that the type of work available in the rural non-agricultural sector might be less desirable than the familiar jobs in agriculture. This instead suggests a role for disamenities of the actual non-agricultural work. Along these lines, we found that rural workers do take non-agricultural jobs when they have fewer opportunities in agriculture. There is a role for future work that further investigates these compensating differentials as a reason why gaps exist between agricultural and non-agricultural wages. Our evidence suggests there may be an equilibrium where the attributes of non-agricultural jobs in rural areas make them less preferred.

References

Alvarez, Jorge. 2018. "The agricultural wage gap: Evidence from Brazilian micro-data." *American Economic Journal: Macroeconomics* forthcoming.

Binswanger-Mkhize, Hans P. 2013. "The stunted structural transformation of the Indian economy." *Economic and Political Weekly* 48 (26-27):5–13.

Bryan, Gharad and Melanie Morten. 2017. "The aggregate productivity effects of internal migration: Evidence from Indonesia." *Journal of Political Economy* forthcoming.

Colmer, Jonathan. 2018. "Weather, labour reallocation, and industrial production: Evidence from India." CEP Discussion Papers (CEPDP1544). Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial wage disparities: Sorting matters!" *Journal of Urban Economics* 63 (2):723–742.

D Narasimha Reddy, N Nagaraj, A Amarender Reddy and Cynthia Bantilan. 2014. "Rural Non-Farm Employment and Rural Transformation in India." *Working Paper Series* (57).

Duncan, Greg J and Bertil Holmlund. 1983. "Was Adam Smith right after all? Another test of the theory of compensating wage differentials." *Journal of Labor Economics* 1 (4):366–379.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell et al. 2015. "The climate hazards infrared precipitation with stations — a new environmental record for monitoring extremes." *Scientific Data* 2:150066.

Gollin, Douglas, Martina Kirchberger, and David Lagakos. 2017. "In Search of a Spatial Equilibrium in the Developing World." National Bureau of Economic Research w23916.

Gollin, Douglas, David Lagakos, and Michael E Waugh. 2014. "The Agricultural Productivity Gap." *Quarterly Journal of Economics* 129 (2):939–993.

Hamory Hicks, Joan, Marieke Kleemans, Nicholas Y. Li, and Edward Miguel. 2017. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata." National Bureau of Economic Research w23253.

Herrendorf, Berthold and Todd Schoellman. 2018. "Wages, human capital, and barriers to structural transformation." *American Economic Journal: Macroeconomics* 10 (2):1–23.

Imbert, Clément and John Papp. 2018. "Costs and Benefits of Rural-Urban Migration: Evidence from India." *Unpublished* .

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E Waugh. 2018. "The Welfare Effects of Encouraging Rural-Urban Migration." National Bureau of Economic Research w24193.

Lagakos, David and Michael E Waugh. 2013. "Selection, Agriculture and Cross-Country Productivity Differences." *American Economic Review* 103 (2):948–980.

Morten, Melanie. 2019. "Temporary migration and endogenous risk sharing in village India." *Journal of Political Economy* 127 (1):1–46.

Munshi, Kaivan and Mark Rosenzweig. 2016. "Networks and misallocation: Insurance, migration, and the rural-urban wage gap." *American Economic Review* 106 (1):46–98.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and aggregate productivity: A quantitative cross-country analysis." *Journal of Monetary Economics* 55 (2):234–250.

Smith, Robert S. 1979. "Compensating wage differentials and public policy: a review." *Industrial and Labor Relations Review* 32 (3):339–352.

Young, Alwyn. 2013. "Inequality, the urban-rural gap, and migration." *The Quarterly Journal of Economics* 128 (4):1727–1785.

Tables

Table 1: Baseline Characteristics

	Ag Only (N=1499)	Switchers (N=387)	p-value
Individual Variables:			
Female	0.388	0.101	0.000***
Years of education	3.477	3.463	0.947
Cognitive ability	2.787	2.708	0.131
Household Variables:			
Household size	5.932	6.214	0.052*
Access to electricity	0.512	0.453	0.038**
House has mud walls	0.674	0.739	0.015**
Number of rooms in house	3.571	3.708	0.169
Area cultivated (acres)	0.575	0.583	0.950
Landless	0.175	0.145	0.159
Has private tubewell	0.038	0.034	0.671
Owns mobile phone	0.933	0.912	0.149
BPL card holder	0.769	0.806	0.122
NREGS job card holder	0.749	0.796	0.053*
NREGS active user	0.193	0.240	0.041**
Scheduled Caste or Tribe	0.517	0.651	0.000***
Has loan	0.167	0.119	0.019**
Has savings account	0.685	0.628	0.032**
Has permanent migrant	0.097	0.098	0.931
Has temporary migrant	0.096	0.140	0.013**

The table shows average values of baseline characteristics between workers that worked only in agriculture for all three surveys that were used to estimate the agricultural wage gap (column 1) and those that worked in both sectors (column 2). Column 3 shows p-value of the t-test for equal means. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Active NREGS user is household that had NREGS income during April 2014, just before the baseline started. Has loan is an indicator for having any loan during the last 12 months. Permanent migrant is individual that is away for at least 10 months of the year. A temporary migrant is defined as an individual that leaves the village during the dry season but returns home during the wet season. Cognitive ability is the score on a reverse digit span test.

Table 2: The agricultural wage gap amongst agricultural laborers

	Individ, Survey (1)	Individ by Survey (2)	Survey (3)	Village, Survey (4)	Village by Survey (5)
Non-ag work	0.207***	0.211**	0.305***	0.325***	0.325***
	(0.041)	(0.083)	(0.040)	(0.036)	(0.035)
Mean ag wages (Rs per day)	169	169	169	169	169
Number workers	2285	2285	2285	2285	2285
Number of Observations	28598	28598	28598	28598	28598
R squared	0.785	0.940	0.315	0.538	0.748

The data are from three surveys where non-agricultural wages were collected: the planting survey of 2015, and the planting and harvesting surveys of 2016. The dependent variable in all columns is the log of daily wages. Column 1 includes individual and survey fixed effects, column 2 includes individual-by-survey fixed effects, column 3 includes only survey fixed effects, column 4 includes village and survey fixed effects, and column 5 includes village-by-survey fixed effects. Columns 1 and 3-5 also include surveyor fixed effects. Observations are weighted by the gender of the respondent, based on the gender shares in the farmers survey. 472 respondents contribute to the identification in column 1 by working in both the agricultural and non-agricultural sectors across the three surveys. 230 workers contribute to the identification in column 2, i.e. they work in both sectors in the same survey round. Standard errors are clustered at the village level in all specifications. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

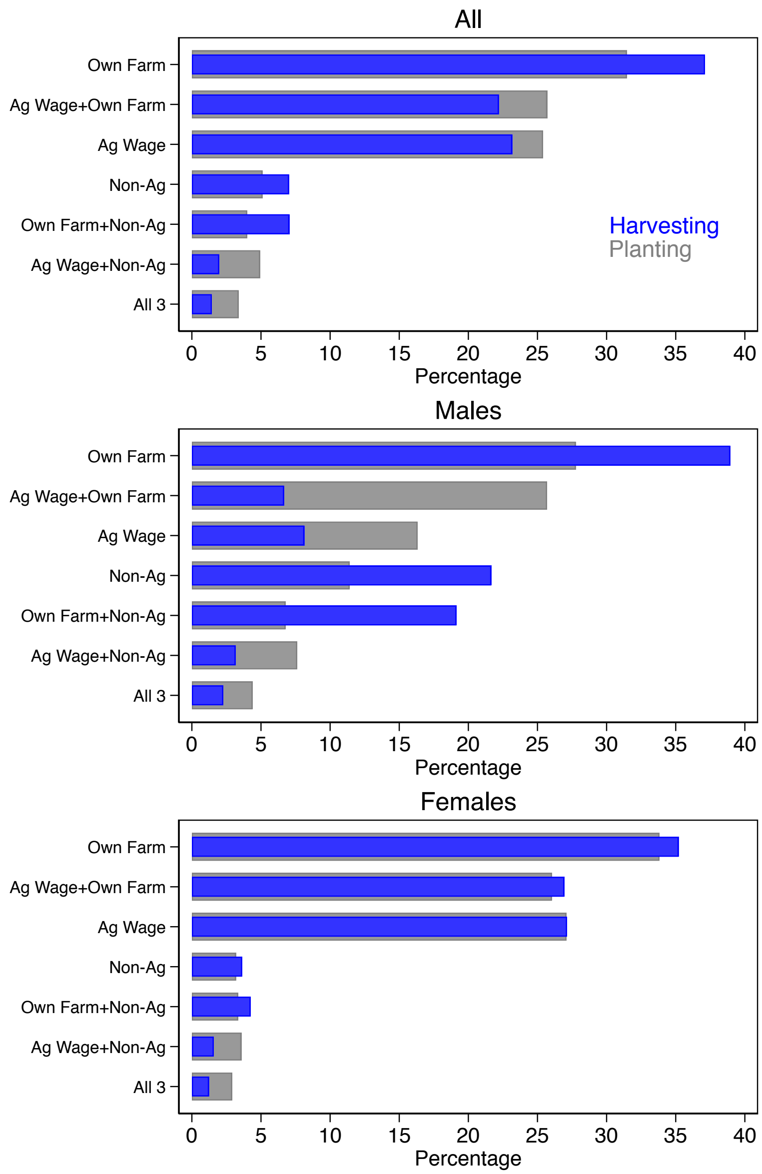
Table 3: Effects of rainfall realizations on agricultural productivity, agricultural earnings of casual laborers, and employment choices

			Daily Activity			
	(1) Log Yield	(2) Ag. Earnings	(3) Ag	(4) Own Field	(5) Non-Ag	(6) Nothing/House
Rainfall	0.520***	8.452***	0.071***	0.036*	-0.051***	-0.045***
June-October	(0.050)	(2.592)	(0.016)	(0.021)	(0.013)	(0.016)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Survey fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.36	34.35	0.22	0.37	0.16	0.24
Number laborers		2645	2645	2645	2645	2645
Number of Observations	5898	78449	78449	78449	78449	78449
R squared	0.463	0.257	0.241	0.140	0.170	0.162

The estimates in column 1 are based on a 3-year panel survey with 2,000 large farmers (10 per village). The dependent variable in column 1 is the log of overall rice yield (across all plots). Columns 2-6 are estimated for the harvesting surveys with agricultural laborers of 2014, 2015, and 2016. The dependent variables are daily earnings from agricultural labor (column 2), an indicator for working in agriculture (column 3), an indicator for working in their own field (column 4), an indicator for working in the non-agricultural sector (column 5), and an indicator for leisure or housework (column 6). The rainfall variable is total rainfall (measured in 100's of mm from June-October). Observations in columns 2-6 are weighted by the gender of the respondent, based on the gender shares in the farmers survey. These regressions also include surveyor fixed effects. Standard errors are clustered at the village level in all specifications. Asterisks indicate a coefficient that is statistically significant at the 1% ***, 5% **, and 10% * levels.

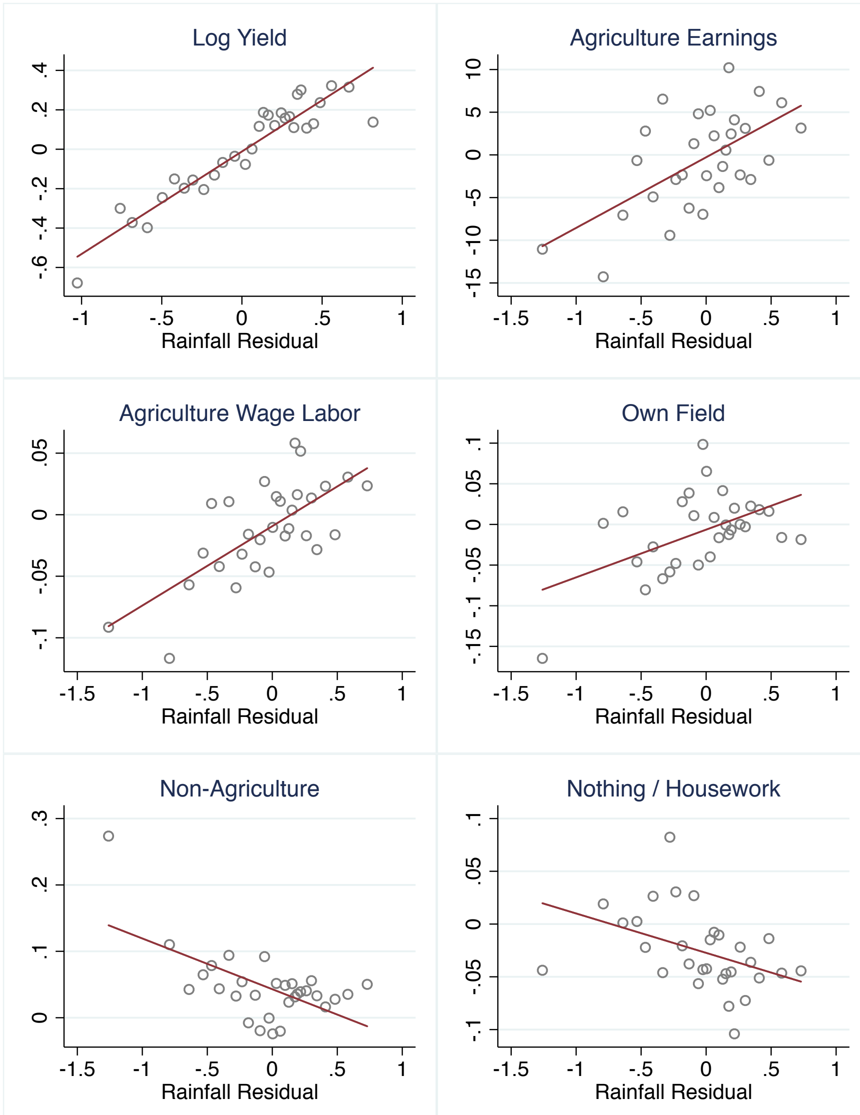
Figures

Figure 1: Activities of workers during 7-14 day survey period



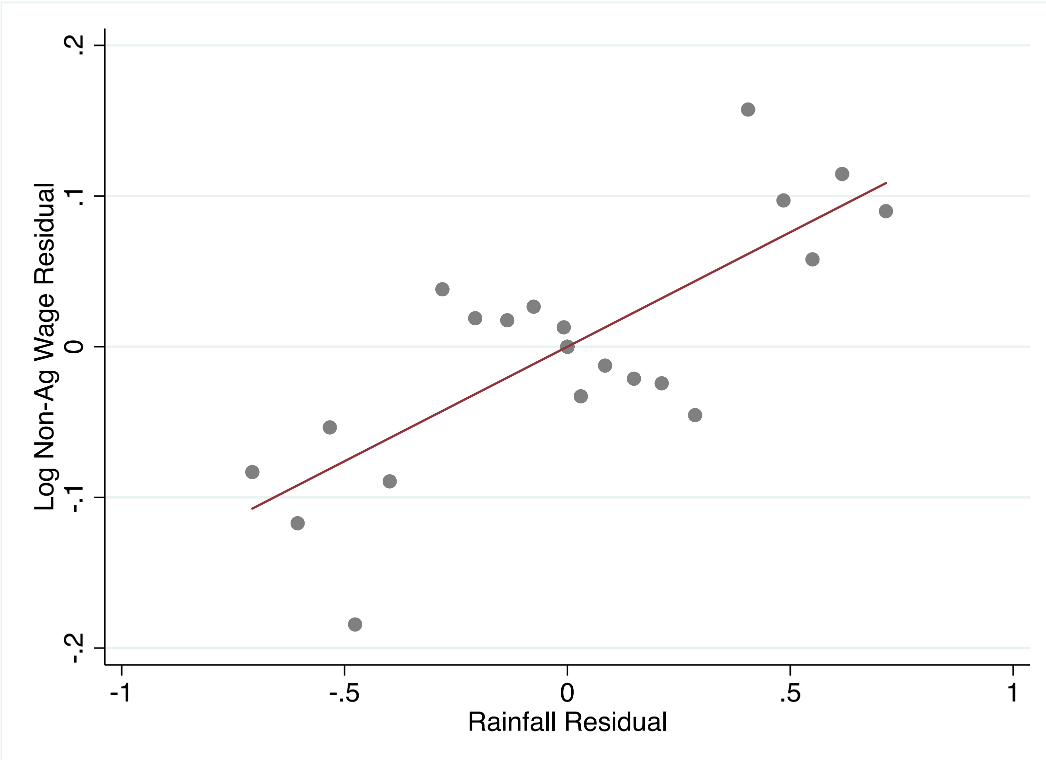
The figure shows a classification of workers into seven groups, depending on which activities they did during the 7 or 14 day survey period. The top panel is for all respondents and is weighted by gender to represent the sex ratios of the population of agricultural workers hired by large farmers. The bottom two panels are separate for males and females. “Own Farm” indicates working on their own farm, “Ag Wage” indicates working for a wage in agriculture, and “non-agricultural” indicates non-agricultural work. The grey bars denote percentages of respondents across the three planting surveys while the blue bars denote the same values for the harvesting surveys. As an example, around 39 percent of the male respondents work only on their own fields during harvesting (top bar in the middle panel).

Figure 2: The relationships between rainfall realizations, agricultural productivity, and labor allocation



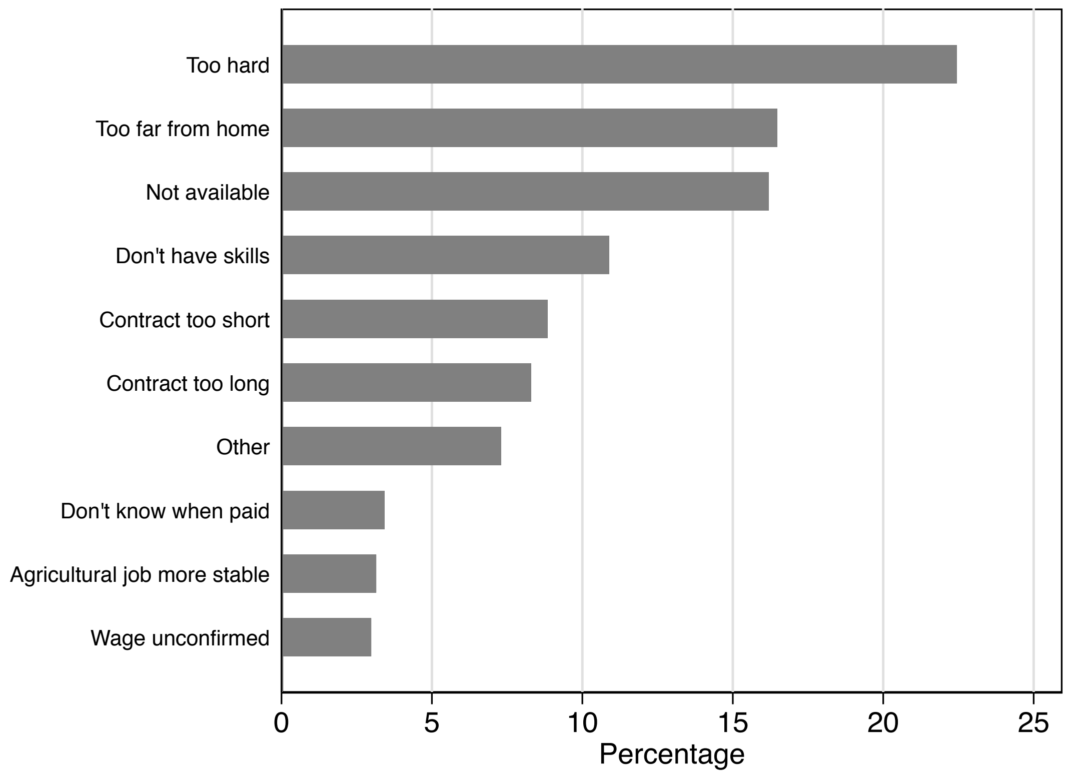
Notes: The graph shows binned scatter plots of various outcomes against rainfall realizations. The data are first residualized by regressing the outcomes and June-October rainfall on surveyor, time, and village fixed effects. Each graph then shows the partial relationship between the outcome and rainfall. The dots are for 30 bins of the rainfall residuals, with equal numbers of observations per bin. The regression line is shown in red. The upper left graph uses the 3-year panel survey with farmers to plot the relationship between rainfall and log rice yield. The remaining outcome variables are from the labor allocation survey with agricultural workers. The outcomes (in order from left to right) are daily agricultural earnings in Rupees, an indicator for working in agriculture as a wage laborer, an indicator for working on the laborer's own field, an indicator for doing non-agricultural work, and an indicator for staying at home or doing housework.

Figure 3: Relationship between non-agricultural wages and rainfall (year 1 and year 3 harvesting)



Notes: The figure shows the relationship between the log of non-agricultural wages and village-level monsoon rainfall. The data are first residualized by regressing the the log of non-agricultural wages and June-October rainfall on village and year fixed effects. The information for year 1 comes from the followup survey, in which a question on non-agricultural wages during harvesting of that year was asked for each household member. The information for year 3 comes from the harvesting phone survey with the sample of laborers. We observe wages for all 200 villages during the year 1 followup survey because we asked about each household member, but we only observed non-agricultural work in 94 unique villages for the year 3 harvesting survey. The regression thus has 294 observations. The coefficient from the regression is 0.15 and the t statistic is 2.21.

Figure 4: Stated reasons why laborers still don't work in the non-agricultural sector even when wages are higher



Notes: The graph shows responses from the third followup survey with agricultural laborers. The exact question posed to laborers was "Suppose wages are a bit lower for agricultural jobs than for non-agricultural jobs, what is the top reason why you may still work in agricultural jobs".

Appendix - for online publication

Table A1: Unweighted estimates of the agricultural wage gap

	Individ, Survey (1)	Individ by Survey (2)	Survey (3)	Village, Survey (4)	Village by Survey (5)
Non-ag work	0.217***	0.175***	0.312***	0.322***	0.325***
	(0.024)	(0.043)	(0.026)	(0.024)	(0.024)
Mean ag wages	169	169	169	169	169
Number workers	2288	2288	2288	2288	2288
Number of Observations	28610	28610	28610	28610	28610
R squared	0.854	0.960	0.485	0.632	0.765

The table presents the same regressions as Table 2 but without weighting observations by gender. The specifications are otherwise the same. The data are from three surveys where non-agricultural wages were collected: the planting survey of 2015, and planting and harvesting surveys of 2016. The dependent variable in all columns is the log of daily wages. Column 1 includes individual and survey fixed effects, column 2 includes individual-by-survey fixed effects, column 3 includes only survey fixed effects, column 4 includes village and survey fixed effects, and column 5 includes village-by-survey fixed effects. Columns 1 and 3-5 also include surveyor fixed effects. Observations are weighted by the gender of the respondent, based on the gender shares in the farmers survey. 472 respondents contribute to the identification in column 1 by working in both the agricultural and non-agricultural sectors across the three surveys. 230 workers contribute to the identification in column 2, i.e. they work in both sectors in the same survey round. Standard errors are clustered at the village level in all specifications. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table A2: Robustness to dropping non-agricultural work outside of the worker's own village

	Individual	Individual by Survey
	(1)	(2)
Non-ag work	0.166***	0.192**
	(0.048)	(0.087)
Mean ag wages	169	169
Number workers	2242	2242
Number of Observations	27236	27236
R squared	0.774	0.936

The data are from three surveys where non-agricultural wages were collected: planting time of 2015, and the planting and harvesting surveys of 2016. This table drops days of non-agricultural work which were classified as outside the village (either migrant labor or when the work was outside the village). The dependent variable in both columns is the log of daily wages. Column 1 includes individual, survey, and surveyor fixed effects. Column 2 includes individual-by-survey fixed effects. Observations are weighted by the gender of the respondent, based on the gender shares in the survey with farmers. 384 respondents contribute to the identification in column 1 by working in both the agricultural and non-agricultural sectors across the three surveys. 205 workers contribute to the identification in column 2, i.e. they work in both sectors in the same survey round. Standard errors are clustered at the village level in all specifications. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table A3: The wage premium for taking agricultural jobs in other villages

	Individ, Survey (1)	Individ by Survey (2)	Survey (3)	Village, Survey (4)	Village by Survey (5)
Work outside the village	0.051**	0.047***	0.036*	0.037**	0.037***
	(0.020)	(0.012)	(0.019)	(0.018)	(0.011)
Mean ag wages	166	166	166	166	166
Number workers	2431	2431	2431	2431	2431
Number of Observations	33533	33533	33533	33533	33533
R squared	0.704	0.993	0.171	0.424	0.807

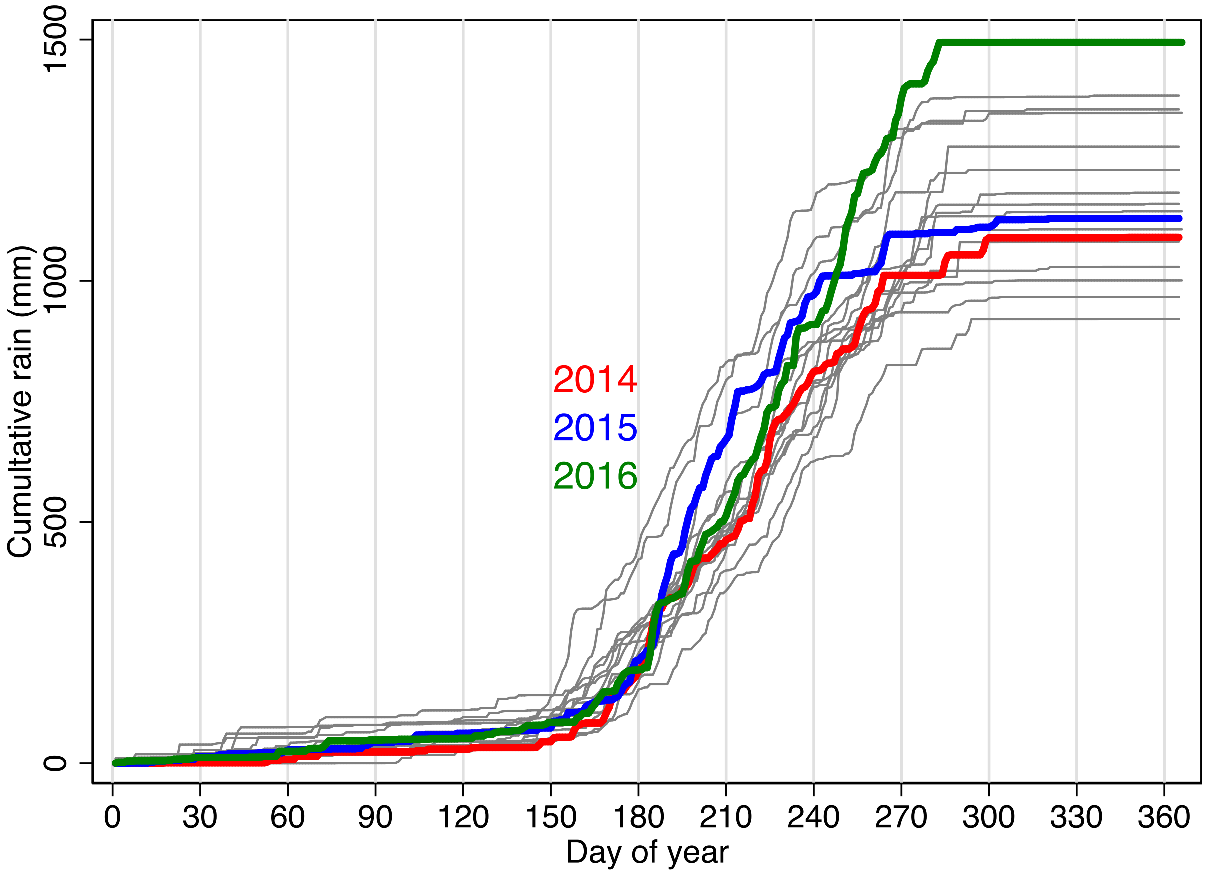
The data are from all six phone surveys and are limited to days when the respondent worked for an agricultural wage. The dependent variable in all columns is the log of daily wages. Column 1 includes individual and survey fixed effects, column 2 includes individual-by-survey fixed effects, column 3 includes only survey fixed effects, column 4 includes village and survey fixed effects, and column 5 includes village-by-survey fixed effects. Columns 1 and 3-5 also include surveyor fixed effects. Observations are weighted by the gender of the respondent, based on the gender shares in the farmers survey. Standard errors are clustered at the village level in all specifications. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Table A4: Correlation between agricultural daily wages and the length of the work day

	Male Log Wages		Female Log Wages	
	(1)	(2)	(3)	(4)
Hours	-0.072***	-0.040*	-0.036***	-0.019*
	(0.019)	(0.021)	(0.013)	(0.011)
Planting	-0.066***	-0.039**	-0.036	-0.083**
	(0.017)	(0.016)	(0.093)	(0.036)
Weeding	-0.094**	-0.036	0.005	-0.064*
	(0.040)	(0.040)	(0.094)	(0.035)
Threshing	-0.014	-0.032***	-0.025	-0.060*
	(0.012)	(0.009)	(0.091)	(0.036)
Harvesting	-0.069**	-0.059***	-0.032	-0.079**
	(0.028)	(0.022)	(0.092)	(0.036)
Village fixed effects	No	Yes	No	Yes
Mean wages (level)	186	186	117	117
Number of Observations	1835	1835	2520	2520
R squared	0.044	0.513	0.013	0.605

The data are from the survey with farmers after the 2014 season. Farmers were asked for male and female wages, separately by task and gender. Farmers were also asked for the length of a typical work day by gender and task. The dependent variables are the log of male wages (columns 1 and 2) and the log of female wages (columns 3 and 4). Standard errors are clustered at the village level in all specifications. Asterisks indicate that coefficient is statistically significant at the 1% ***, 5% **, and 10% * levels.

Figure A1: Cumulative rainfall in study area, 2000-2016



The figure shows cumulative rainfall plotted against the day of the year. Each line is for a separate year. Daily rainfall was first averaged across the 200 sample villages to generate a daily average precipitation for the sample area. The daily rainfall values are satellite observations taken from CHIRPS.