GLM|LIC
Growth and Labour Markets in Low Income Countries Programme

WP

GLM|LIC Working Paper No. 31 | June 2017

How Complete Are Labor Markets in East Africa?

Evidence from Panel Data in Four Countries

Brian Dillon (University of Washington)
Peter Brummund (University of Alabama)
Germano Mwabu (University of Nairobi)

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid from the British people

GLM|LIC

GLM|LIC Working Paper No. 31 | June 2017

How Complete Are Labor Markets in East Africa?

Evidence from Panel Data in Four Countries

Brian Dillon (University of Washington)
Peter Brummund (University of Alabama)
Germano Mwabu (University of Nairobi)

GLM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Fax: +49-228-3894-510
Email: glm-review@iza.org

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid from the British people

GLM|LIC Working Paper No. 31 | June 2017

ABSTRACT

How Complete Are Labor Markets in East Africa?

Evidence from Panel Data in Four Countries*

We develop new tests for the completeness of rural labor markets. The tests are based on a theoretical link between a shortage or surplus in the labor market and asymmetric responses to changes in household composition over time. We develop auxiliary tests to distinguish other types of market failures from labor market failures, and provide evidence that most changes in household composition are exogenous to local labor market conditions. We implement our test using nationally representative panel data from Ethiopia, Malawi, Tanzania, and Uganda. The overall pattern is one of excess supply of labor in rural areas, but there is substantial heterogeneity across cultivation phases, genders, and agro-ecological zones. Excess supply of labor is most evident during low-intensity cultivation phases (e.g., weeding). In Ethiopia, findings suggest that poor households face a *de facto* labor shortage, driven more by financial market failures than a physical shortage of available workers. There is evidence of partial gender segmentation in labor markets. In all four countries, women are more difficult to replace than men.

JEL Classification:

O13, O15, J20, J43

Keywords:

agricultural households, labor markets, separation, asymmetric adjustment, East Africa

Corresponding author:

Peter Brummund
Department of Economics, Finance and Legal Studies
University of Alabama
P.O. Box 870224
Tuscaloosa, AL 35487
USA
E-mail: pbrummund@cba.ua.edu

* We thank the Institute for the Study of Labor (IZA) and the UK Department for International Development for funding through the GLM-LIC program. Preliminary work on this project took place as part of the Myths and Facts in African Agriculture project of the World Bank and African Development Bank. For comments and helpful discussions we are grateful to Chris Barrett, Taryn Dinkelman, Dave Donaldson, An- drew Foster, Louise Fox, Markus Goldstein, Doug Gollin, Rachel Heath, Joe Kaboski, Supreet Kaur, David Lam, Simone Schaner, Tavneet Suri, and seminar participants at the Pacific Development Conference, CSAE Conference in Oxford, and numerous GLM-LIC conferences organized by IZA. We thank Joshua Merfeld, Reuben Mutegi, and Audrey Royston for excellent research assistance. All errors are the responsibility of the authors.

1 Introduction

Labor is the primary endowment of the world's poor households. Poverty reduction is directly linked to increases in the returns to labor, whether through higher wages in the market or higher productivity in self-employment. As a key input to agriculture, labor also contributes to the majority of output in rural areas of low-income countries. For these reasons, improving our understanding of how labor markets function in rural areas has been central to the research and policymaking agenda since at least Lewis (1954) and Harris and Todaro (1970).

In this paper we develop a new test to detect non-clearing labor markets, and to identify whether such markets are in excess supply or excess demand. Our approach builds on a classic literature in development economics that uses the resource allocation problem of family farms to test the completeness of rural markets (Sen, 1966; Singh et al., 1986; Benjamin, 1992; Barrett, 1996; Udry, 1999; Le, 2010; Dillon and Barrett, 2017; LaFave and Thomas, 2016). The idea behind the classic test is as follows. In the standard model of the agricultural household, the profit-maximization problem of the farm is embedded in the household's utility maximization problem. If markets for inputs, outputs, and other relevant goods are complete and competitive, the household's consumption and production problems are separable. The family farm can be analyzed as a profit-maximizing firm, and household endowments have no impact on the input demand functions of the farm. Hence, testing whether there is a relationship between household endowments and farm inputs is tantamount to testing the complete markets assumption.

The most common way to implement this *separation test* is to regress farm labor utilization on the household labor endowment. A positive and significant relationship indicates that households with more (fewer) members use more (less) labor on their farms – a violation of separation, and hence of the assumption of complete markets. In a seminal paper, Benjamin (1992) shows in a cross-section from Indonesia that farm labor demand does not depend on the number of workers in the household, leading to non-rejection of the complete markets assumption. In a recent advance, LaFave and Thomas (2016) come to the opposite conclusion using panel data from the same setting. In both both pooled and fixed effects regressions, LaFave and Thomas find a significant relationship between labor endowments and labor utilization on farm, leading them to reject the complete markets assumption.

A limitation of the standard test is that a rejection of separation does not identify the specific pattern of underlying market failures. Although the test is implemented using data on household labor endowments and farm labor utilization, separation can fail even if labor markets are complete (Feder, 1985; Udry, 1999). Interpretation is especially difficult in a cross section, where time invariant factors such as heterogeneity in managerial skill or preferences for working on one's own farm can lead to non-separation. Hence, while the standard test provides insights into the completeness of markets, and is informative for model selection, its practical use as a guide for policymakers is limited.

We show that in panel data, a variation of the separation test can sometimes provide insights into the state of the rural labor market, specifically. The intuition is as follows. Suppose that in period $t-1$, a household faces a binding ration on the number of hours it can work in the market (a labor demand constraint), perhaps because of a downward sticky wage. Such rationing can lead to non-separation, with household members working on the family farm up to a point at which the marginal revenue product of farm labor is below the market wage. Now suppose that from period $t-1$ to period t, someone exits the household. If the reduction in the household's labor endowment relieves the binding ration on market work, separation becomes possible. Farm labor falls, but only to its optimal level. The opposite is not true: if the household labor endowment increases from period $t-1$ to period t, the ration continues to bind, non-separation persists in period t, and farm labor increases. The implication is that in a large sample, a binding labor demand constraint predicts a specific pattern of *asymmetric* average responses to increases and decreases in labor endowments. These predictions apply even if the household enjoys separation in period $t-1$.

A binding ration on the supply of labor – i.e., a lack of available workers – predicts the opposite set of asymmetries. In this case, farm labor utilization responds more to decreases in labor endowments than to increases. Hence, by testing for asymmetries in the average response of farm labor usage to increases and decreases in the household labor endowment, we can test necessary conditions for binding constraints on labor demand and labor supply. This test requires panel data, because identification is from within-household changes in labor endowments over time.

In Section 2 we formally develop this test, and consider a number of potentially confounding issues. We first determine whether failures in the markets for credit, land, insurance, or other inputs lead to similar asymmetric predictions. For those that do, we develop additional tests that distinguish a non-clearing labor market from other possible causes of asymmetric non-separation. We also consider whether changes in labor endowments might be endogenous to labor market conditions, with households explicitly recruiting or releasing members in response to their labor market experience. If changes in labor endowments are endogenous to local conditions, an excellent instrument is the average change in labor endowment of households in the same village, excluding one's own change. Ultimately, we show that this instrument is very weak, in all study countries. We interpret this as evidence that changes in labor endowments are largely exogenous, driven by factors such as marriage, divorce, death, boarding school, and children aging into the workforce. Descriptive patterns in the data, as well as the findings from qualitative work that we conducted to augment this paper, support this interpretation. We explain how endogenous adjustments, if present, will only attenuate the asymmetries that are the basis of our identification strategy.

Section 3 provides information on the data and sample. We implement our tests using the nationally representative Living Standard Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) data from four East African countries: Ethiopia, Malawi, Tanzania, and Uganda. Because the data sets are national in scope, our findings provide a characterization of the average state of rural labor markets in each country. To test whether labor market segmentation might lead to variation in labor market conditions for specific subgroups, we conduct additional tests for heterogeneity by the gender composition of the labor endowment, and by agro-ecological zone. To the extent permitted by the data, we also allow for variation in labor market conditions across cultivation phases (planting, weeding, harvest), out of concern that short-term spikes in labor demand for particular activities could lead to different labor market conditions at different times.

Section 4 presents the empirical findings. In regressions that impose symmetry, separation is rejected in all study countries. The estimated elasticity of farm labor utilization to the household labor endowment ranges from 0.55-0.65 for Ethiopia, Malawi, and Tanzania, and is roughly half that magnitude in Uganda. When we allow for asymmetric non-separation and variation across cultivation phases, some intriguing differences emerge across countries. Findings for Malawi and Uganda are clearly consistent with a binding ration on off-farm work, i.e., a general pattern of excess labor supply in rural areas. Results for Tanzania lean in the same direction, although the asymmetry is less pronounced. In Ethiopia we find the opposite: the evidence is consistent with a binding labor supply constraint. We find some important level differences across cultivation phases, consistent with long periods of under-utilized labor between peaks of more efficient resource allocation, and across agro-ecological zones. Finally, we find evidence of partial gender segmentation in labor markets. Labor supply constraints are more likely to bind for women, and labor demand constraints are more likely to bind for men. That is, labor supplied to the farm by female household members is less likely to be replaced in the market than that supplied by their male counterparts.

In Section 5 we discuss the findings and dig deeper into the puzzling results for Ethiopia. A point of emphasis in the discussion is that asymmetric non-separation does not rule out failures in other (non-labor) markets, but it does reveal whether a non-clearing labor market is a factor preventing separation. We test the hypothesis that the Productive Safety Net Programme, a large-scale workfare program in Ethiopia, might be crowding out labor supply to the private market. There is little evidence to support this hypothesis. However, we do find that the labor supply constraint in Ethiopia only binds for poor households; non-poor households exhibit asymmetric responses consistent with a binding demand constraint (i.e., a lack of off-farm opportunities). This implies that asymmetric non-separation in Ethiopia is due to a mix of factors: financial market constraints prevent poor households from farming at optimal intensity, while labor demand constraints prevent non-poor households from working the desired number of hours in the market.

We view the contribution of this paper as threefold. First, we provide a generalization of the standard separation test, one that allows us to learn about the average condition of local labor markets (when we find certain patterns of results). Other papers employing this class of tests, with the notable exception of Udry (1999), focus exclusively on testing the completeness of rural markets.

Second, we provide evidence consistent with a prevailing pattern of excess supply of rural labor, for all but the poorest half of households in Ethiopia. This aligns with other evidence on nominal wage rigidities (Dreze and Mukherjee, 1989; Osmani, 1990; Kaur, 2016) and on rural-urban productivity gaps (Gollin, Lagakos and Waugh, 2014; McCullough, 2017), which can be interpreted as broad evidence of over-supply of rural labor. Collectively these findings underscore the lack of non-farm opportunities for households in rural areas of East Africa. We further show that non-separation is more extreme during less intense cultivation periods. This implicates the technology of non-mechanized farming as a factor in the incompleteness of labor markets. The need to provide substantially more labor during brief but critical periods leads to an over-supply of rural labor during other times of year.

Finally, the findings directly demonstrate the importance of catering research and policy to specific local conditions, even in neighboring countries from a single region. In this introduction we have provided a general overview of results. Yet, this overview masks substantial heterogeneity. In the body of the paper we provide more details about the heterogeneity in both symmetric and asymmetric non-separation across countries, cultivation phases, genders, and agro-ecological zones. The clear lesson is that field data from a single trial in any one area of the study countries would not provide a representative picture of underlying market conditions for the region as a whole.

2 Theory and empirical framework

We begin this section by developing a dynamic version of the standard agricultural household model (Sen, 1966; Singh et al., 1986; Benjamin, 1992). Our model is similar to that in LaFave and Thomas (2016). In Section 2.2 we develop the tests that associate specific labor market failures with asymmetric responses to changes in labor endowments, and in Section 2.3 we consider other types of market failures. Section 2.4 deals with identification, and Section 2.5 describes additional dimensions of heterogeneity that we test later in the paper.

2.1 A dynamic agricultural household model

Consider a farming household endowed with E_t units of labor in year t. The household divides its labor endowment between leisure L_t^l, work on the household farm L_t^h, and supply of labor to the market, L_t^m. The household has preferences over consumption C_t and leisure L_t^l, represented by the strictly increasing and concave utility function $U(C_t, L_t^l)$. The household farm produces the single consumption good C using strictly increasing, concave production technology $F(L_t)$, where L_t represents total labor application, and other inputs are subsumed in the production function (we discuss one key input, land, in Section 2.4). Total output is $y = F(L_t)\epsilon_t$, where ϵ_t is an exogenous production shock representing the multiplicative effects of various sources of uncertainty over the value of output, including those due to weather, pest pressure, or output prices. The household can hire labor on the market, represented by L_t^d. Let w_t be the market wage rate, with the price of the output normalized to 1.

The household has access to credit markets in which it can borrow or lend at interest rate r_t. Hence, liquid wealth W_{t+1} is equal to $1 + r_t$ times the difference between wealth at the start of period t, W_t, and net income in period t.

If markets are complete and competitive, and utility is inter-temporally separable, then the household's utility maximization problem takes the following form:

$$\max \quad \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t U(C_t, L_t^l \mid \epsilon_t, \rho_t)\right] \tag{1}$$

$$\text{subject to:} \quad C_t - w_t L_t^m + \frac{1}{1+r}(W_{t+1} - W_t) \leq F(L_t)\epsilon_t - w_t L_t^d \tag{2}$$

$$L_t = L_t^h + L_t^d \tag{3}$$

$$E_t = L_t^h + L_t^m + L_t^l \tag{4}$$

$$L_t^l, L_t^h, L_t^d, L_t^m, C_t, \geq 0 \tag{5}$$

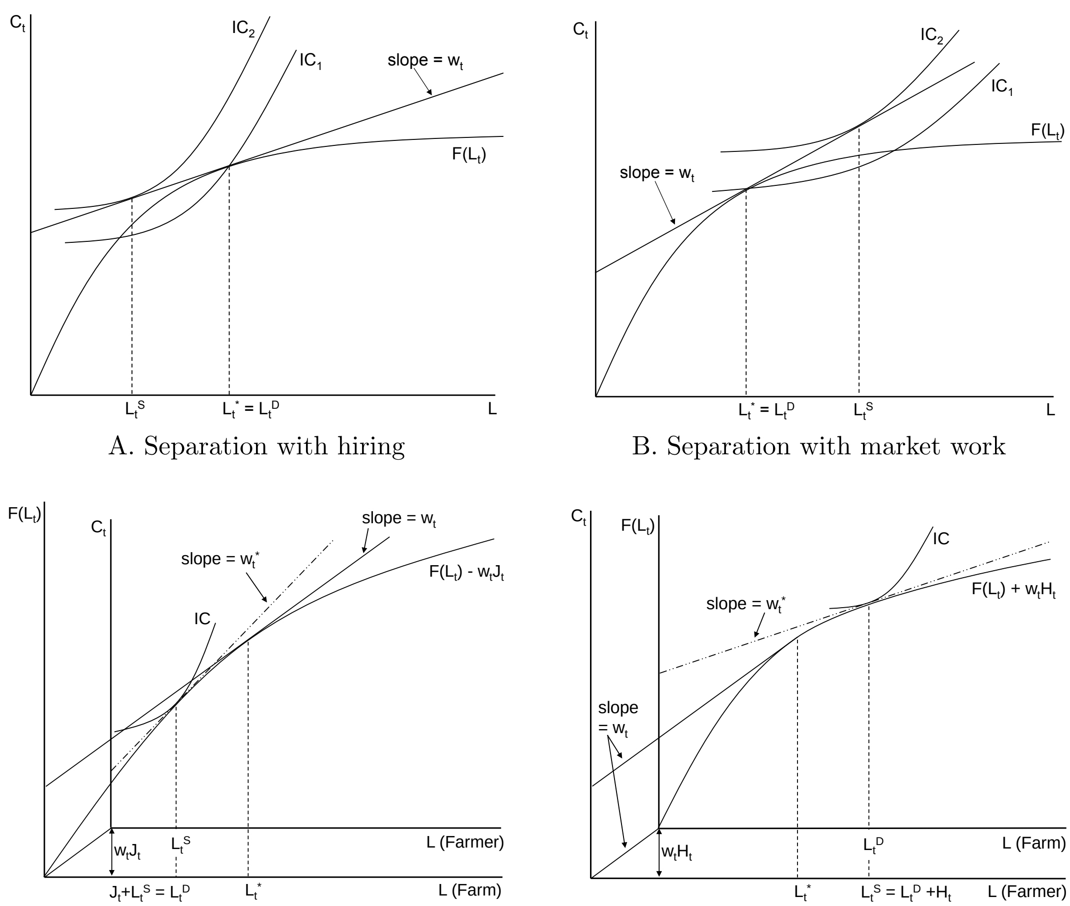
where the utility function is conditioned on the stochastic output shock and a parameter ρ_t that represents preferences and endowments. The equality in (2) will hold at the solution. Under current assumptions, the model is recursive, and the consumption and production sides of the household problem can be solved separately. In period t, household members first choose L_t to maximize expected farm profit, which is on the right-hand side of (2). They then maximize utility, conditional on expected farm income. The solution is characterized by the following:

$$\Pi^* = \max \mathbb{E}\left[F(L_t^*)\epsilon_t - w_t L_t^{d^*}\right] \tag{6}$$

$$L_t^D \equiv L_t^* = L_t^{h*} + L_t^{d*} = L^D(w_t, r_t \,|\, \epsilon_t) \tag{7}$$

$$L_t^S \equiv L_t^{h*} + L_t^{m*} = L^S(w_t, r_t \,|\, \rho_t) \tag{8}$$

where equation (6) is the expected profit function, equation (7) is the farm labor demand function, and (8) is the household labor supply function. The complete markets assumption imposes the testable exclusion restriction that labor demand is not a function of the household labor endowment, E_t, which is a component of ρ_t. With complete markets, L_t^* depends only on prices and ϵ_t. If all markets but one are complete, then separation holds, because relative prices can adjust to accommodate one non-tradable (Feder, 1985).



A. Separation with hiring

B. Separation with market work

C. Non-separation with supply constraint

D. Non-separation with demand constraint

Figure 1: Household labor supply and farm labor demand

Notes: With only minor adaptations: Panel B is based on Figure 1 in Benjamin (1992); Panel C is based on Figure 3 in the same paper; Panel D is based on Figure 2 in the same paper. Panel A is original to this paper.

In Figure 1, panels A and B show two possible cases under separation.[1] In panel A, the marginal rate of substitution at L_t^*, the point at which the marginal revenue product of labor on farm is equal to the wage, is such that the household would rather work less than L_t^*. This household provides L_t^S units of labor to the farm, and the farm hires additional workers up to the point L_t^*. Panel B shows the opposite case. The household prefers to work beyond the point L_t^*, hence it provides additional work to the market after supplying the optimal amount of labor to its own farm.

This characterization of separation suggests that a household would not both buy and sell labor in the same season. In fact, we can allow for that possibility with only minor adjustments. One option would be to add to the model a utility value from sometimes working on others' farms, possibly related to learning, socializing, or maintaining group cohesion to support coinsurance. An empirically tractable and plausible alternative would be to allow for variation across phases of production. A household's position in the labor market might be different during periods of peak demand (planting, harvest) than during other cultivation periods, if wages and the shadow value of farm labor vary between periods.

With access to panel data, we can test the exclusion restriction implied by equation (4) while controlling for household fixed effects (LaFave and Thomas, 2016). If non-separation in a cross-section is due to time invariant (or very slowly evolving) factors, such as managerial skill, preferences for working on one's own farm, or household size preferences determined through fertility choices, then we should find that violations of separation in the cross-section disappear in the panel.

2.2 Incomplete labor markets and asymmetric non-separation

In this subsection we build on the dynamic model of the previous subsection to consider the implications of an incomplete labor market for household labor supply and farm labor demand. To fix ideas, assume that the labor market and some other market are incomplete, so that separation does not hold. Benjamin (1992) extensively develops the theory governing household labor allocation decisions under non-separation for the two types of labor market failures of primary interest in this section. The first scenario is one in which the household

[1] The graphs in Figure 1 are based on figures in Benjamin (1992), with only minor modifications.

faces a ration H_t on the number of hours it can provide to the market, perhaps because of a downwardly sticky wage. The second is one in which the household farm faces a hard limit, J_t, on the labor that it can hire in the market, perhaps because local markets do not adjust quickly enough to spikes in demand related to certain activities.

Panels C and D of Figure 1 show the implications of these two different scenarios for household labor supply and farm labor demand. In panel C, preferences are such that the household does not want to provide the additional labor required to reach L_t^*, on top of the maximum amount that it can hire from the market (J_t). The household optimizes by providing only L_t^S to its own farm, and the shadow value of labor is w_t^*, which is above the market wage. In panel D, preferences are such that even after providing H_t labor to the market and working on the family farm, household members would prefer to work more. They cannot do so in the market, so instead they supply additional labor to the farm, up to the point L_t^S. At this optimum, w_t^* is below w_t.

Benjamin (1992) shows that under either of these types of non-separation, the critical question is: How does the shadow wage vary with the household labor endowment? That is, what is the sign of dw_t^*/dE_t in the cross-section? He further shows that under the plausible assumption that equilibrium labor supply increases with the labor endowment – i.e., the addition of a household member does not raise the utility value of leisure so substantially that *total* household labor supply falls, hence $dL_t^S/dE_t > 0$ – then $dw_t^*/dE_t < 0$ in either case. This leads to the testable prediction that labor utilization on farm is increasing in the household labor endowment under either of these types of non-separation.

In a static model, these predictions are based on infinitesimal changes in labor endowments, evaluated through differentials. The key comparative static, dL_t^S/dE_t, is equivalent to dL_t^D/dE_t, and both are symmetric, by construction. In a dynamic setting, however, households experience discrete changes in their labor endowments from one period to the next. This introduces the possibility of asymmetric responses to changes in endowments, and the asymmetries differ depending on the type of labor market failure.

To see this, suppose first that a non-separating household in period $t-1$ experiences an increase in labor endowment from period $t-1$ to period t, i.e., $\Delta E_t > 0$. Under the analogous assumption to Benjamin, that total desired household labor supply cannot fall when the labor endowment increases, this has the effect in panels C and D of Figure 1 of tilting the indifference curve (IC) to the right (in either case). The marginal rate of substitution at the previous optimum is now lower, and the new optimum moves to the right, with increases in labor supply by the household, labor demand on the farm, and consumption. In the panel D case, this discrete shift serves only to exacerbate the effect of the ration H_t, and L_t^D increases 1-for-1 with L_t^S. However, if the household is initially like the one in panel C, then the increase in E_t may relieve the worker shortage, so that the supply constraint no longer binds. In that case, the household sets $L_t^D = L_t^*$, and supplies any additional labor to the market. The increase ΔE_t has made separation possible, and the new equilibrium looks like that in panel B.

By an analogous line of reasoning, when the labor endowment falls from one year to the next ($\Delta E_t < 0$), L_t^D and L_t^S for a supply-constrained household (panel C) fall together. But for a household facing labor demand constraint H_t in period $t-1$ (panel D), the reduction in the labor endowment may be such that the constraint no longer binds and separation becomes possible. For this household, the new optimum is like that in panel A.

Without knowing the exact shapes of the production function and the preference map, we cannot make predictions about the relative magnitudes associated with these changes. But if preferences and production are smooth in the neighborhood of the initial optimum, then in a large sample of households facing supply constraint J_t we expect to see a larger average response of L_t^D to endowment decreases ($\Delta E_t < 0$) than to endowment increases ($\Delta E_t > 0$), because some of the increases in L_t^D will be truncated at L_t^*. Indeed, if L_{t-1}^D is "close" to L_{t-1}^* for many households like that in panel C, then the response of ΔL_t^D to a discrete increase in ΔE_t is unlikely to be statistically different from zero. Similarly, if L_{t-1}^D is "close" to L_{t-1}^* for many households like that in panel D, then the response of ΔL_t^D to a decrease in ΔE_t will be difficult to detect.

To test these predictions, we estimate the following empirical models using nationally representative panel data from the four study countries (separately):[2]

[2]The theory also suggests a test comparing the relative magnitudes of ΔL_t^S and ΔL_t^D to positive and negative changes in endowments. Such tests are not possible in the LSMS-ISA data, because the survey module on household labor supply is incompatible with that on farm labor demand.

$$
\begin{aligned}
L_{ht}^{D} &= \beta_0 + \beta_1 E_{ht} + \gamma A_{ht} + \gamma_2 demog_{ht} + \nu_t + \mathbf{EA} + \epsilon_{fht} && (9)\\
\Delta L_{ht}^{D} &= \beta_0 + \beta_1 \Delta E_{ht} + \gamma \Delta A_{ht} + \gamma_2 \Delta demog_{ht} + \nu_t + \epsilon_{fht} && (10)\\
\Delta L_{fht}^{D} &= \mathbf{F} + \beta_1 \Delta E_{ht} + \beta_2 \mathbf{F} \Delta E_{ht} + \gamma \Delta A_{ht} + \gamma_2 \Delta demog_{ht} + \nu_t + \epsilon_{fht} && (11)\\
\Delta L_{fht}^{D} &= \mathbf{F} + \beta_1 \Delta E_{ht}^{+} + \beta_2 \Delta E_{ht}^{-} + \beta_3 \mathbf{F} \Delta E_{ht}^{+} + \beta_4 \mathbf{F} \Delta E_{ht}^{-} \\
&\quad + \gamma_1 \Delta A_{ht} + \gamma_2 \Delta demog_{ht} + \nu_t + \epsilon_{fht} && (12)
\end{aligned}
$$

where t indexes year, h indexes household, and f indexes cultivation phase (planting, weeding, harvest); quantity of labor demanded (L_{fht}^{D}), labor endowment (E_{ht}), and cultivated acreage (A_{ht}) are entered in logs; the $^{+}$ and $^{-}$ superscripts on ΔE_{ht} indicate increases and decreases, respectively; $\mathbf{F}$ is a vector of dummy variables for cultivation phases; *demog* is a vector of controls for the demographic breakdown of the household; ν represent time effects interpretable as levels in equation (9) and differences in equations (10)–(12); $\mathbf{EA}$ are fixed effects for enumeration areas (roughly equivalent to villages); and ϵ is a statistical error term.

Equation (9) is a pooled model. Equations (10) and (11) are household fixed effects model in first differences, similar to the main specifications in LaFave and Thomas (2016). In models (9) and (10) we aggregate all farming activities into a single measure of total farm labor demand, while (11) allows for heterogeneity across cultivation phases. Specifications (9)–(11) provide a basis of comparison for equation (12), a household fixed effects model that allows for asymmetric responses of farm labor utilization to increases and decreases in the household labor endowment, and allows for variation across cultivation phases. Equation (12) is the main specification of interest.

In all models, we expect the β_k, $k = 1, \ldots, 4$, to be non-negative, and non-separation is consistent with $\beta_k > 0$. The prediction of the theory in this subsection is that when the average household faces the binding labor demand constraint H_t in the cultivation phase excluded from $\mathbf{F}$, β_1 estimated in (12) will be of greater magnitude and statistically different from zero with greater probability than β_2. When the average household faces the binding labor supply constraint J_t during the excluded cultivation phase, we expect the opposite: β_2 will be greater in magnitude and more likely to be statistically different from zero than β_1.

In the other cultivation phases represented by $\mathbf{F}$, the relevant comparisons are between the total marginal effects, $\beta_1 + \beta_3$ for increases and $\beta_2 + \beta_4$ for decreases.

Of course, in large, nationally representative data sets, households in one area may face supply constraints while those in another are face demand constraints. The power of the test to detect asymmetries clearly weakens as the proportion of households facing each type of constraint becomes roughly similar. The test is also weaker if there is widespread non-separation, symmetric or otherwise, due to other types of market failures. Finally, the labor market failures considered so far are sufficient for asymmetric non-separation, but not necessary. Hence, we next consider whether other types of market failures lead to similar asymmetric predictions.

2.3 Failures in other markets

It is well understood that non-separation identified in the labor dimension can stem from various underlying patterns of market failures, even when labor markets are complete. In this subsection we discuss those possibilities and consider whether other patterns of market failures could generate asymmetric predictions like those derived above.

2.3.1 Credit

Suppose first that the labor market is complete, but the markets for credit (and some other good) are not. A separating household like that in panel A of Figure 1 uses its liquid resources to hire L_t^{d*} units of labor in the market. However, without a complete credit market, the shadow value of cash to the household may be above the market interest rate. In that case the household cannot hire L_t^{d*} and achieve $L_t^D = L_t^*$. This household effectively faces a labor supply constraint, and looks like the non-separating household in panel C of Figure 1, even though labor markets are complete. This prediction is not symmetrical: there is no general reason that a household for whom the shadow value of cash is below the market interest rate should over-supply labor to its own farm.

The implication is that even with complete labor markets we might still see the asymmetry associated with binding supply constraint J_t: significant changes in L_t^D when

ΔE_t is negative, but smaller or statistically insignificant changes when ΔE_t is positive. This is a potential challenge to the interpretation of the test proposed in Section 2.2.

If we see this pattern, to distinguish a labor supply constraint from a credit constraint we re-estimate (12), interacting the terms involving $\mathbf{F}$ and ΔE_{ht} with a binary variable that takes a value of 1 if the household is above median wealth. If we find that households across the wealth spectrum exhibit asymmetric responses consistent with a labor supply constraint, that is an indication that the shortcoming is in the labor market. If the asymmetric response is concentrated among households that are more likely to be credit-constrained, that suggests that the main driver is an incomplete credit market.

Note that this test does not tell us whether the entire structure of the rural economy, including the demand for labor outside of agriculture, is distorted by a lack of access to finance. Identification here is based on within-household variation in farm labor utilization as a function of the household labor endowment, and hence can only provide insights into whether asymmetric non-separation, specifically, is driven by incomplete access to credit for the farming households in the data.

2.3.2 Land

Dillon and Barrett (2017) use the same data employed here to show that land markets are clearly functioning in many areas of sub-Saharan Africa. Yet, the existence of a rental market, sales market, or other institution for temporary assignment of land use rights does not guarantee that such markets are complete or competitive. Incomplete land markets can clearly introduce rigidities that lead to non-separation. The specific concern is that an inability to adjust the amount of land under cultivation might prevent households from optimally responding to changes in the labor endowment. Furthermore, there is a potential asymmetry in how incomplete land markets impact separation. It is always possible to reduce the amount of land under cultivation, by leaving some of it fallow,[3] but incomplete land markets might make it difficult to increase cultivated land from one period to the next.

Following Benjamin (1992) and LaFave and Thomas (2016), we control for culti-

[3]Though fallowing may be unwise in some settings, if it reduces future claims on the plot (Goldstein and Udry, 2008).

vated land in specifications (9)–(12). Hence, our results capture the relationship between the change in the household labor endowment and the change in labor-used-per-acre. The rationale for including a control for land, but not other physical inputs, is that land can be treated as fixed once the cultivation season begins. The concern with that, which is related to the concern in the previous paragraph, is that households might anticipate problems finding workers when deciding how much land to plant, which introduces a form of simultaneity into the land and labor decisions.

To address these concerns, we re-estimate (9)–(12) using acres owned as an instrument for acres cultivated. While households do buy and sell agricultural land in some areas, sales are not as common as adjustment of cultivated acreage through rental markets, borrowing for free, or fallowing. Hence, if the relationships between ΔL_t^D, ΔE_t^+ and ΔE_t^- are driven by incomplete land markets, they should disappear or be severely attenuated in the IV.

2.3.3 Insurance

Suppose now that the labor market is complete, but the market for insurance is not (and there is at least one other missing or incomplete market). In this case the household bears all risk associated with uncertainty in the final value of output, represented by ϵ_t in (1). It is straightforward to see that the lack of an insurance market can lead to non-separation and a correlation between labor usage on farm and the labor endowment (Udry, 1999). The intuition comes from examining the first order condition for farm labor, L_t, under the assumption that labor markets are complete but the household bears all output risk. Letting $g(\epsilon_t)$ represent the density function of ϵ_t, the first order condition for L_t^* is:

$$U_{C_t}(C_t^*, L_t^{l*}) \left[\int_{\epsilon_t} F'(L_t^*)\epsilon_t g(\epsilon_t) - w_t \right] = 0 \tag{13}$$

where U_{C_t} is the derivative of the utility function with respect to consumption. If the household labor endowment increases from one period to the next, the optimal demand for labor on farm, L_t^*, need not be directly affected, because the labor market is complete. But optimal leisure, L_t^{l*}, changes with the increase in the labor endowment, and this impacts the marginal utility of consumption (the first term in (13)), which depends on ϵ_t through the first order condition for C_t. Hence, the value of L_t^* that solves (13) changes with the labor endowment, even though the labor market is complete, because the household cannot sell its production risk and thereby predict the marginal value of consumption with certainty.

In this case there is no *a priori* reason to expect asymmetric adjustment to ΔE_t^+ and ΔE_t^-. As long as the utility function is smooth in the neighborhood of observed consumption and leisure, households respond to changes in either direction. There are no thresholds beyond which $L_t^* \equiv L_t^D$ ceases to adjust, as in the labor market case from Section 2.2. For this reason, no additional test is needed to distinguish insurance market failures from labor market failures. But if non-separation is driven by incomplete insurance markets, this will raise the share of household exhibiting symmetric responses, and thereby decrease the power of the test based on asymmetric responses.

2.3.4 Any two other non-labor goods

Finally, suppose that the labor market is complete, but the markets for any two other relevant goods are not complete. This is akin to a standard general equilibrium model with at least two non-tradables. In such a situation, relative prices cannot adjust to clear markets. Hence, households could face supply constraint J_t or demand constraint H_t simply because markets are not clearing generally, rather than because of a specific feature of the labor market.

In one sense, this situation is not problematic for the empirical strategy in Section 2.2. The labor market is still not clearing, and the pattern of asymmetric responses to changes in labor endowments can still potentially reveal whether the average household faces a labor supply or labor demand constraint. Such an insight is important and policy relevant regardless of whether a specific feature of the labor market is the fundamental cause of non-clearing. However, the possibility that a non-clearing labor market could be a side effect of some other shortcoming serves as a cautionary note for the interpretation of results. The asymmetric test can (sometimes) provide evidence about whether the average household is supply- or demand-constrained, and that in itself is an important input to policymaking, but the test cannot distinguish between possible root causes of incomplete labor markets.

2.4 Identification

In Section 2.1 we noted that many potential causes of non-separation are time invariant (or slowly evolving), e.g., preferences for working on one's own farm, or managerial skill. The inclusion of household fixed effects in models (11) and (12) controls for any such factors. Household fixed effects also control for any long-term household planning, e.g., a household that wanted to farm intensively might have had more children in anticipation of possible market failures. Furthermore, in Section 2.3 we developed tests that under certain conditions can distinguish labor market failures from failures in other markets.

Despite these steps, one key identification concern remains. We have implicitly treated changes in the household labor endowment as exogenous to the labor market. Many such changes are clearly exogenous: household composition evolves with marriage, divorce, illness, death, children growing up, the beginning or ending of boarding school, and various other factors. However, at least in theory, a household facing a labor supply or labor demand constraint in period $t-1$ might endogenously adjust its period t labor endowment by recruiting new household members, or sending some away. Such adjustments clearly reflect labor market failures: a wholesale change in the composition of the household is an extreme response to a labor allocation problem, and is unnecessary if local labor markets are working well. But if large numbers of households endogenously adjust their labor endowments in this manner, the power of the asymmetric test developed in Section 2.2 decreases. In the extreme, if *all* adjustments are endogenous, then the predicted asymmetry is the opposite of that in Section 2.2: a period t significant relationship between decreases in labor endowment and decreases in labor utilization on farm would be simultaneous responses to binding demand constraint H_{t-1}, whereas our current prediction is that ΔE^-_{ht} leads to greater magnitude adjustments than ΔE^+_{ht} when the household faces supply constraint J_t.

If changes in labor endowments are endogenous to labor market conditions, then households facing the same labor market will tend to exhibit correlated changes in endowments. This suggests that an excellent instrument for ΔE_{ht} is $\Delta \overline{E}_{et}^{-h}$, the mean change in labor endowments for households in enumeration area e (which contains household h), excluding the change of household h itself. Enumeration areas are roughly equivalent to villages across the study countries. Hence, a strong first stage indicates substantial within-village correlation in ΔE_{ht}, in which case we cannot rule out that the primary source of identifying variation is from households endogenously responding to local conditions. Conversely, a weak first stage would indicate that in our samples, changes in endowments are largely exogenous to local labor market conditions.

In Appendix B we provide first stage F-statistics from these IV regressions. Across all study countries, the instrument is very weak. F-statistics in the main specification are all below 4, and half are below 2. Results do not improve with alternative specifications (varying how we count children in the labor endowment; excluding enumeration areas with few households). There is only one F-statistic across all variants that is above 5, and it is 6.5. Taken together, this set of findings is a strong rejection of the hypothesis that variation in labor endowments is due to households adjusting to local labor market conditions.

This finding corroborates other evidence, from descriptive statistics and from focus groups that we conducted in study countries, suggesting that most of the variation in labor endowments is driven by factors other than the local labor market. Figure S1 in Appendix B shows scatter plots of $\Delta\overline{E}_{et}$ across countries and waves. If changes were correlated locally, we would expect to see most points clustered along the two axes. Instead, changes are distributed throughout the simplex with no clear pattern. In 2015 we conducted a series of focus groups in Malawi and Tanzania to discuss labor allocation problems broadly. In every focus group, respondents roundly rejected the idea that households recruit or send away members based on their inability to find farm labor or off-farm jobs.

Respondents' clear rejection of the idea of endogenous adjustments may overstate the case. For instance, a household might be more willing to release a young adult to migrate if that person is not needed on the farm and cannot find work locally, but may not in conversation characterize that decision as a response to local labor market conditions. Nevertheless, the first stage IV results corroborate our focus group findings, and suggest that factors other than local labor constraints are the primary drivers of changes in household composition. Hence, our main results are based on OLS rather than IV specifications. Endogenous adjustments, if present, will only attenuate the asymmetric effects predicted by the theory in Section 2.

2.5 Additional dimensions of heterogeneity

Finally, given that non-clearing labor markets are the focus of the paper, a natural extension to our main analysis is to test for variation in asymmetric non-separation across potentially segmented labor markets. We test for such heterogeneity along two dimensions. The first is gender. Specifically, we use the gender identify of each household member to define separate labor endowments E_{male} and E_{female}, and re-estimate the main specification (12). These tests are motivated by a large body of research on gender-differentiated outcomes in agriculture (Besley, 1995; Udry, 1996; Yngstrom, 2002; Allendorf, 2007; Goldstein and Udry, 2008; Ubink and Quan, 2008; Kumar and Quisumbing, 2012; Ali, Deininger and Goldstein, 2014; Doss et al., 2015; Dillon and Voena, 2016). If men and women are not perfect substitutes, due to discrimination or to differences in the shadow value of male and female labor to the household, then we may find different results across gender lines.

The second dimension of heterogeneity is by agro-ecological zone (AEZ). Differences in the agro-climate – including rainfall patterns, the range of average temperatures, and the degree of humidity – may impact the labor market in numerous ways, for example by changing the duration of each cultivation phase, altering the mix of planted crops (which could influence the number of required weedings or the time sensitivity of planting), or changing the number of crop cycles in the year. For each country we identify the 2-3 most common AEZs and estimate (12) with interactions of all $\mathbf{F}$ and ΔE terms with dummy variables for AEZ.

3 Data, sample, and descriptive patterns

We test the predictions of the above model using panel data from the Living Standards Measurement Study and Integrated Surveys on Agriculture (LSMS-ISA). These are comprehensive household and agricultural surveys, conducted by national statistics offices with cooperation from the World Bank. The data are nationally representative, span a wide range of topics, and are reasonably comparable across countries. The four study countries are those for which at least two waves of LSMS-ISA panel data were available when we conducted the analysis. We use two waves of panel data from Ethiopia, two from Malawi, three from Tanzania, and four from Uganda, covering the following time periods: Ethiopia, 2011-2012 and 2013-2014; Malawi, 2010-2011 and 2013; Tanzania, 2008-2009, 2010-2011, and 2012-2013; Uganda, 2005-2006, 2009-2010, 2010-2011, and 2011-2012. More details about the survey activities for each country are provided in Appendix A.

Because our analysis uses the labor demand equation of the household farm, we restrict the samples to households that report cultivation of a positive number of acres in more than half of survey waves. For Ethiopia and Malawi this excludes all households that did not cultivate in both survey years. For Tanzania and Uganda, this excludes households that did not cultivate in more than one survey year.

LSMS-ISA questionnaires are based on a standard template, and hence are broadly similar across countries. Each survey begins with a household roster asking for the names of individuals who normally live and eat their meals together. We use the list of such members to construct the variable for the labor endowment of household h in period t, E_{ht}, by counting the number of working age household members. Later in this section we provide a rationale for our definition of "working age."

The agriculture survey modules share some similarities across countries. Every survey distinguishes between household labor and non-household labor. In most cases there is more detail about household labor than hired labor.[4] Every survey also allows for plot-specific reporting of some variables. However, there is substantial variation between countries in how plots are defined, and in some cases there is no way to link plots across survey waves. Hence, we aggregate the agricultural variables (acreage and labor demanded) to the household level. The agricultural surveys also differ in the degree of disaggregation by labor activity. Questionnaires for Ethiopia and Malawi distinguish between non-harvest and harvest activities, which we use to form variables for labor demand during "Cultivation" and "Harvest." The Tanzania data are even further disaggregated into "Planting", which includes land preparation and planting, "Weeding", which includes weeding and applying top-dressing

[4]For example, the household labor module may ask the total number of weeks worked, average number of days per week, and average number of hours per day for each person who worked on the plot, while the hired labor module may ask only about the total number of men from outside the household who worked on the plot and the average number of days a man was hired, and then repeat the those questions for women and children. In Appendix A we provide details about the construction of the farm labor demand variable L^D_{ht} for each country.

fertilizer, and "Harvest." The Uganda survey does not differentiate between activities, hence we can only construct a single variable for "All farming activities" in Uganda.[5]

Table 1: Summary statistics

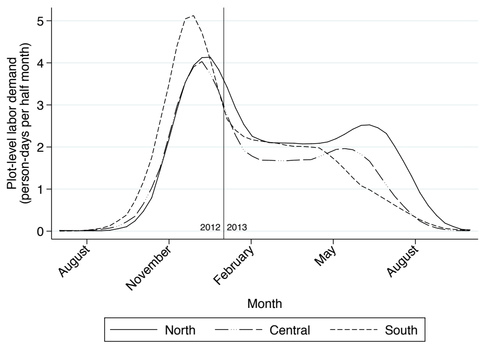
	Ethiopia		Malawi		Tanzania		Uganda	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Labor endowment, with kids	3.03	1.42	3.06	1.51	3.38	1.82	3.30	1.77
Labor endowment, no kids	2.72	1.28	2.78	1.37	3.06	1.67	2.91	1.58
Prime male share	0.23	0.17	0.23	0.17	0.24	0.19	0.23	0.19
Prime female share	0.26	0.17	0.25	0.16	0.26	0.16	0.24	0.16
Elderly male share	0.08	0.13	0.03	0.12	0.04	0.13	0.04	0.12
Elderly female share	0.07	0.15	0.05	0.15	0.06	0.17	0.04	0.14
Acres cultivated	5.1	22.4	2.0	1.7	5.9	11.7	4.2	10.8
Acres owned	3.8	16.7	1.8	1.7	5.4	11.7	3.5	9.7
Age of head (years)	45.5	15.0	44.5	16.2	49.8	15.5	46.6	15.2
Education of head (years)	1.6	2.8	5.6	4.3	5.1	3.3	4.5	3.3
Reference labor (person-days)	146.0	447.3	76.6	77.6	61.7	74.6	136.0	137.8
Harvest labor (person-days)	83.1	150.2	22.6	36.5	49.0	74.1		
Weeding labor (person-days)					56.3	63.6		
Number of Obs.	5676		4818		6062		8266	

Notes: Authors' calculations from LSMS-ISA data. "Labor endowment with kids" uses adult equivalence scale for children aged 11-15, defined later in this section. "Prime" demographic groups are those aged 15-60; "Elderly" are aged 61+. The "Reference" agricultural phase is "All non-harvest" for Ethiopia and Malawi, "Land preparation and planting" for Tanzania, and "All farming activities" for Uganda.

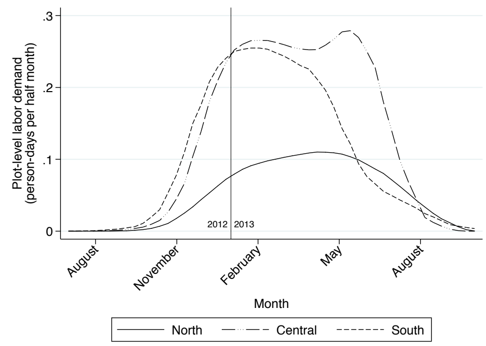
Table 1 provides relevant summary statistics for each country, pooled across survey waves. In the top two rows we see that the average household has a labor endowment of roughly 3 working age adults, with significant variation (below, we describe two ways of calculating the labor endowment, with and without phasing in children as they age). The demographic composition variables, for the number of prime-age (15–60) and senior (60+) household members by gender, are included in all empirical specifications. Households in Malawi have substantially smaller farms than those in other countries. Throughout the paper we use "Reference" to refer to the first cultivation phase in each country, so that we can collect results into multi-country tables. There is wide variation across countries in total labor application. However, because the structure of the agricultural labor survey

[5] In a small number of cases, respondents report zero labor for one activity but a positive amount of labor for other activities. These zeroes may be measurement error, but they may also have an economic rationale, e.g., it is possible to apply zero weeding labor, or to apply zero harvest labor if the crop fails. We assign these zeroes a small, positive value, so that they are not dropped when we take logs. In our main tables this value is 0.1 person-days. All of our findings are robust to other reasonable replacement values, and to dropping these observations entirely.

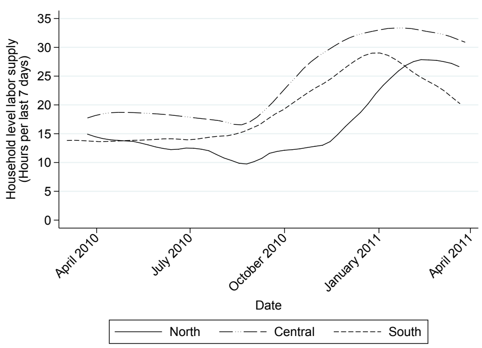
varies across countries, we cannot disentangle real differences, due for instance to variation in the length of the growing season, from those due to framing or measurement error. In this respect it is useful that identification is based on within-household changes over time.



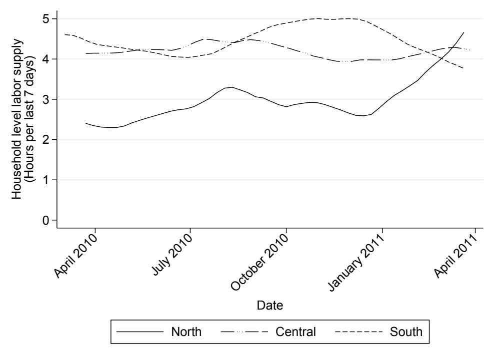
A. Farm use of household labor



B. Farm use of hired labor



C. Household labor supply to own farm



D. Household labor supply to local market

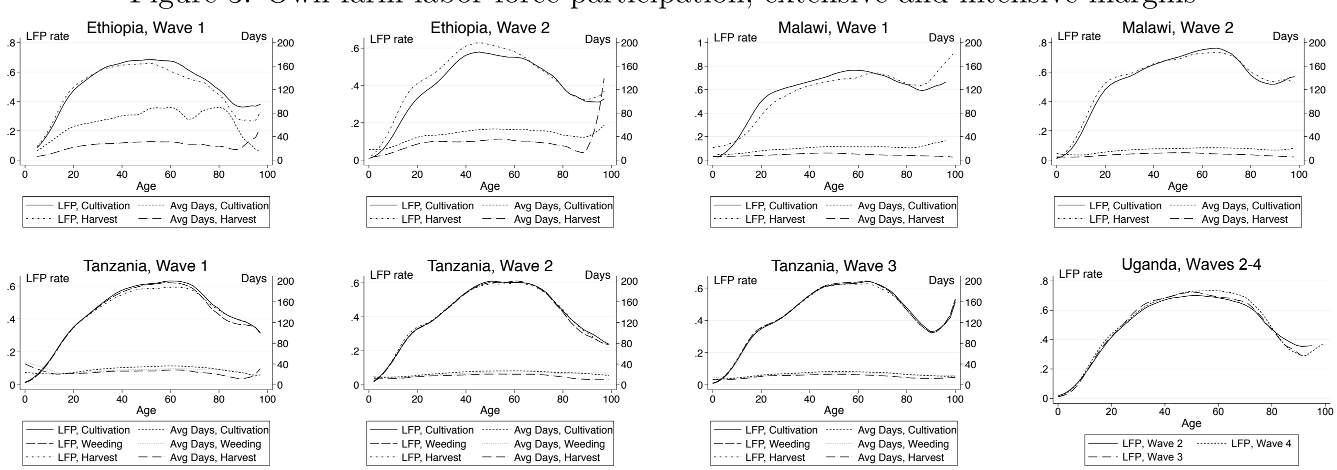
Figure 2: Time path of labor demanded and supplied, Malawi

Notes: Authors' calculations from LSMS-ISA data. Top panel is from the long rainy season in the 2012-2013 survey; bottom panel is from 2010-2011 survey. All graphs are local polynomial regression using an Epanechnikov kernel. Additional details for panel A are in Appendix A.3.

In Section 2 we discussed the possibility of spatially correlated spikes in the demand for labor during peak periods. Figure 2 shows kernel regressions of labor demanded on farms (panels A and B) and labor supplied by households (panels C and D) across time, separately for the three regions of Malawi. The top two panels are based on the 2012-2013 agriculture module from Malawi, which is the only data set in which we can match labor activities to specific dates. There are three takeaways from panels A and B. The first is that the amount of household labor dwarfs that of hired labor (compare the vertical axes). The second is that the largest spike in labor demanded is associated with planting, which occurs toward the end of the calendar year. Planting begins earlier in the South region than in the Central or the North, which matches the timing of the onset of the rains. The third takeaway is that labor utilization increases around harvest (May-July), but the spike is not as pronounced as that at planting, and is not present in all regions. This may reflect the fact that for some crops, farmers have more leeway with the timing of the harvest than they do with planting. However, it may also reflect covariant production shocks leading to low yields that year.

The lower panels of Figure 2 are based on the 2010-2011 household labor module from Malawi, which is the only data set collected over a full calendar year. Panels C and D show kernel regressions of the time spent working on one's own farm and time spent supplying *ganyu* labor (casual farm labor) over the last seven days (note that the units, hours per last 7 days, are different from those of the top panels). Once again we see that labor supply to own farms is much greater than that to the market, at all times of year. Furthermore, intra-annual variation in labor supply to the market is less pronounced than that to own farms. The latter increases by 100-300% from peak to trough, while the former ranges from near-constant to at most a 50% increase from peak to trough. The general pattern of intra-annual dynamics matches the labor demand side, with the peaks in own-farm labor supply occurring first in the South, then the Central region, then the North.

Figure 3: Own-farm labor force participation, extensive and intensive margins



Authors' calculations from LSMS-ISA data. Uganda data do not include a breakdown by activity and do not allow for differentiation in work days at the individual level.

We have alluded on multiple occasions to changes in labor endowments that occur through the aging of household members. To define E_{ht} we must choose age cutoffs at which someone enters or exits the labor endowment. To guide this decision, we examine own-farm labor supply, by age, for each data set. Figure 3 shows kernel regressions of the own-farm labor force participation rate (LFP), the extensive margin, and the average number of days worked by those who are working, the intensive margin, plotted against age. Separate plots are shown for each cultivation phase. In all figures the axes are scaled so that the LFP lines appear above those for days worked.

Older people do significant work on farms. In all panels of Figure 3, the LFP rate for 70-year-olds is higher than that for 30-year-olds, and the rate for 80-year-olds is higher than that for 20-year olds. These ranges cover most of the senior population (only a fraction of a percent of sample members is over age 80).[6] In general, the drop-off in LFP between ages 60 and 80 is more gradual than the increase during youth. The most rapid changes in own-farm LFP occur between ages 10 and 20. There is also little variation by age in the average days worked, conditional on working. Across countries and activities, 40-year-olds and 70-year-olds work roughly the same number of days. Finally, there is little meaningful variation between farming activities (by age), on either margin.

Based on these observations, we count all adult household members, including senior citizens, in our definition of E_{ht}. At the other end of the age distribution, we allow children to gradually age into the workforce with a linear adult equivalence scale from age 11 onwards: 11-year-olds count as 0.2 adults in the labor endowment, 12-year-olds as 0.4, and so on to age 15.[7] As a robustness check, we also use a binary cut-off at age 15. The first two rows of Table 1 refer to these two methods of including children in the labor endowment, with "no kids" referring to the binary cut-off at age 15.

With this definition of the labor endowment, ΔE_{ht} can be non-zero for three reasons: new people move in, previous household members move out (or pass away), or children age into the workforce. Table 2 shows summary statistics for these changes. All entries are household-level means, with children aged 11-15 scaled in the manner described above (except for the first row, which counts all household members as one person). The second row gives the average net change in the size of the labor endowment, which is then decomposed

[6] The upper tails of the age distributions are as follows: ET wave 1, 2.0% are over age 70, 0.6% over age 80; ET wave 2, 2.1% over 70 and 0.5% over 80; MW wave 1, 2.0% and 0.6%; MW wave 2, 1.9% and 0.6%; TZ wave 1, 2.6% and 0.6%; TZ wave 2, 2.7% and 0.9%; TZ wave 3, 2.8% and 1.0%.

[7] The ages of children working on farm were not recorded for Ethiopia, hired labor in Malawi, hired labor in waves 2 and 3 for Tanzania, and Uganda. In these cases, we count each child worker as 0.5 adults.

Table 2: Inter-annual changes in number of members and labor endowment

	Ethiopia	Malawi	Tanzania		Uganda		
Change between waves:	1 & 2	1 & 2	1 & 2	2 & 3	1 & 2	2 & 3	3 & 4
Δ Number of members	0.04	0.44	0.35	-0.04	0.40	-0.31	0.03
Δ Labor endowment (E)	0.05	0.35	0.25	0.04	0.28	-0.15	0.05
Δ E from move-ins	0.21	0.36	0.31	0.27	0.76	0.21	0.30
Δ E from move-outs	-0.43	-0.41	-0.34	-0.52	-1.04	-0.56	-0.46
Δ E from aging children	0.27	0.40	0.28	0.29	0.56	0.19	0.20
Any net Δ in E (=1)	0.69	0.71	0.68	0.71	0.80	0.71	0.71
Increase in E (=1)	0.43	0.53	0.50	0.45	0.51	0.41	0.47
Decrease in E (=1)	0.26	0.18	0.18	0.26	0.30	0.31	0.24

Notes: Authors' calculations from LSMS-ISA data. Entries are household-level averages. Children between ages 11 and 15 are counted as (Age-10)*0.2 in the labor endowment. Δ Labor endowment is the sum of the three categories immediately below.

into move-ins, move-outs (which includes deaths), and aging into the workforce. The final three rows show the proportion of households experiencing any change in labor endowment, a positive change, or a negative change, respectively. The most important takeaway is that approximately 70-80% of households experience a net change in labor endowment from one survey to the next. Hence, the majority of surveyed households contribute to identify the effects of interest. The average reduction from move-outs is greater in magnitude than the average increase from move-ins, but the average net change in labor endowment is positive in all but one survey wave, after accounting for aging children. Net increases are roughly twice as common as net decreases.[8]

4 Results

In this section we present the main empirical results. The first subsection reports the main findings, followed by robustness checks and analysis of heterogeneity.

4.1 Main results

Table 3 reports the marginal effects of interest from estimates of equations (9)–(12), separately for the four study countries. In the top panel, columns 1, 3, 5, and 7 report the

[8] See Appendix section C for details on the activities performed by in-migrants, out-migrants, and stayers. There is no indication that households recruit new members for their farming skill.

coefficient on E_{ht} in from pooled model (9). There we see that separation is strongly rejected in all four countries. The elasticity of farm labor to the labor endowment is 0.52-0.65 for Ethiopia, Malawi, and Tanzania, and 0.31 for Uganda.

Columns 2, 4, 6, and 8 of the top panel of Table 3 report the coefficient on ΔE_{ht} from model (10), the fixed effects model with farm labor aggregated across cultivation activities. All elasticities remain highly statistically significant. For Ethiopia, Malawi, and Tanzania, the panel result is greater, but similar in magnitude to the pooled result. In those countries, time invariant household factors are not responsible for non-separation in the pooled model. In Uganda, the elasticity falls to 0.20 when we include household fixed effects, a reduction of a third. This indicates that fixed household characteristics are important, but not exclusively responsible, for non-separation in Uganda.

The magnitudes of the estimated elasticities provide a rough guide to the degree of non-separation, i.e., of the extent to which households cannot use local markets to solve their resource allocation problems. Evidence in the top panel of Table 3 suggests that the market distortions facing households in Ethiopia, Malawi, and Tanzania have a greater impact on resource allocation than do those facing households in Uganda.

Columns 1, 3, 5, and 7 in the lower panel of Table 3 show estimates of model (11), household fixed effects models that allow for heterogeneity across cultivation activities. The column 7 result for Uganda is identical to column 8 in the top panel, because the Uganda survey does not differentiate between cultivation activities. For the other three countries, allowing for variation across cultivation phases does not change the main result: non-separation holds in every country, in every phase. There are some level differences in the estimated elasticities, most notably in Ethiopia, where the coefficients on ΔE are roughly 25-30% smaller than in column 2 of the top panel. Perhaps surprisingly, for all three countries the estimated coefficients are smallest in magnitude (though still large) during the harvest period. F-tests reported below the coefficients indicate that the harvest coefficient is significantly different from the reference (pre-harvest) coefficient, in Malawi, and the weeding coefficient, in Tanzania.

Finally, columns 2, 4, 6, and 8 in the lower panel of Table 3 show estimates of model (12), the main specification of interest, which allows for differences across activities and

Table 3: Testing for symmetric and asymmetric non-separation in panel data

Dependent variable:	Log of farm labor (person-days), for columns 1, 3, 5, 7 in top panel Δ Log of farm labor (person-days), all other models							
	Ethiopia		Malawi		Tanzania		Uganda	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All activities								
Labor endow. (E)	0.581***		0.520***		0.652***		0.310***	
	(0.000)		(0.000)		(0.000)		(0.000)	
ΔE		0.623***		0.557***		0.710***		0.204***
		(0.000)		(0.000)		(0.000)		(0.002)
Observations	5,652	2,825	4,818	2,095	5,901	3,895	8,218	5,362
R^2	0.558	0.088	0.432	0.103	0.252	0.025	0.275	0.090
By activity								
Reference $\times \Delta E$	0.471***		0.623***		0.662***		0.204***	
	(0.005)		(0.000)		(0.000)		(0.002)	
Reference $\times \Delta E^+$		0.179		0.869***		0.682***		0.267***
		(0.462)		(0.000)		(0.002)		(0.009)
Reference $\times \Delta E^-$		0.697***		0.236		0.645***		0.147
		(0.006)		(0.342)		(0.010)		(0.114)
Harvest $\times \Delta E$	0.444***		0.321**		0.607***			
	(0.008)		(0.015)		(0.000)			
Harvest $\times \Delta E^+$		0.177		0.389**		0.774***		
		(0.495)		(0.043)		(0.000)		
Harvest $\times \Delta E^-$		0.649**		0.216		0.447*		
		(0.019)		(0.426)		(0.066)		
Weeding $\times \Delta E$					0.784***			
					(0.000)			
Weeding $\times \Delta E^+$						0.940***		
						(0.000)		
Weeding $\times \Delta E^-$						0.635***		
						(0.010)		
Reference: Inc = Dec	.	0.170	.	0.073	.	0.919	.	0.416
Harvest: Inc = Dec	.	0.267	.	0.647	.	0.356	.	.
Weeding: Inc = Dec	.	.	.	.	.	0.393	.	.
Reference = Harvest	0.792	.	0.001	.	0.479	.	.	.
Reference = Weeding	.	.	.	.	0.104	.	.	.
Weeding = Harvest	.	.	.	.	0.012	.	.	.
Reference = Harvest, +	.	0.992	.	0.001	.	0.489	.	.
Reference = Harvest, −	.	0.765	.	0.871	.	0.107	.	.
Reference = Weeding, +	.	.	.	.	.	0.036	.	.
Reference = Weeding, −	.	.	.	.	.	0.940	.	.
Weeding = Harvest, +	.	.	.	.	.	0.117	.	.
Weeding = Harvest, −	.	.	.	.	.	0.057	.	.
Observations	5,650	5,650	4,190	4,190	11,685	11,685	5,362	5,362
R^2	0.028	0.028	0.152	0.153	0.028	0.028	0.090	0.090

Notes: Authors' calculations from LSMS-ISA data. Top panel, odd-numbered columns, are pooled regressions, with enumeration area fixed effects. Bottom panel and top panel, even-numbered columns, are household fixed effects regressions implemented in first differences. Estimates are marginal effects, interpretable as elasticities. p-values in parentheses. Significance: *** 0.01, ** 0.05, * 0.1. Standard errors clustered at the enumeration area. All regressions include controls for changes in log of cultivated acreage, changes in demographic shares, and year fixed effects. The "Reference" activity is "All non-harvest" for Ethiopia and Malawi, "Planting and land preparation" for Tanzania, and "All farming activities" for Uganda. Labor endowment LE is measured in logs in all specifications. ΔE^+ is the change in labor endowment if positive, and 0 otherwise; ΔE^- is the change in labor endowment if negative, and 0 otherwise. Both the dependent variable and the labor endowment are defined so that everyone over age 15 is 1 worker, and children aged 11-14 count as $(Age - 10) \times 0.2$ workers.

asymmetric responses to increases and decreases in the labor endowment. Some important differences now emerge across the study countries.

In column 2 we see that non-separation in Ethiopia is driven primarily by decreases in the labor endowment. The coefficients on "Reference $\times \Delta E^{-}$" and "Harvest $\times \Delta E^{-}$" are 0.7 and 0.65, and highly statistically significant. The coefficients on the ΔE^{+} variables, in contrast, are both equal to 0.18, with p-values approaching 0.5. The imprecision of the ΔE^{+} estimates is such that we cannot reject equality of the increase and decrease coefficients at standard levels – p-values are 0.17 and 0.27. Nevertheless, this pattern is exactly in line with the predictions of the model when labor supply constraint J_t is binding for the average household: decreases in the labor endowment lead to changes in farm labor that are larger in magnitude and much more likely to be statistically different from zero than increases in the labor endowment. The implication is that the average household in Ethiopia is faced with an insufficient supply of workers, in both cultivation phases.

This is a surprising finding, and it contradicts other recent evidence on labor market performance in Ethiopia. However, as described in Section 2.3, this pattern of responses is necessary, but not sufficient, for a binding labor supply constraint. In the next section we dig deeper into this result for Ethiopia.

Results for Malawi show the opposite pattern. In column 4 of the lower panel of Table 3, the only statistically significant relationships between farm labor and the labor endowment are from *increases* in the endowment. The asymmetry is most pronounced in the pre-harvest (reference) period, when the point estimates differ by 0.636, and the p-value for an F-test of symmetry is 0.073. Yet the general pattern holds in the harvest period, as well, with a coefficient on ΔE^{+} that is larger in magnitude and more likely to be different from zero than the coefficient for ΔE^{-}. F-tests reported in the lowest part of the table show that the difference between the "Reference $\times \Delta E^{+}$" and "Harvest $\times \Delta E^{+}$" coefficients is highly statistically significant. Taken together, the results for Malawi are consistent with a binding labor demand constraint H_t – a lack of off-farm opportunities – that has a greater effect on household labor allocation during the pre-harvest period than during the harvest.

Estimates for Tanzania are in column 6 of the lower panel. All of the estimated coefficients are positive and statistically significant, and range from 0.45–0.94. In all three cultivation phases, the elasticity for ΔE^+ is greater in magnitude and has a lower p-value than that for ΔE^-. This is suggestive of a binding labor demand constraint, particularly in the harvest and weeding periods when the magnitude of the difference is substantial. However, the asymmetry is less pronounced than in Ethiopia and Malawi. We interpret this is at most suggestive evidence of this specific labor market shortcoming. Interestingly, farm labor during the weeding period exhibits the greatest response to ΔE^+. Off-farm opportunities for household members are particularly scarce during this period.

Results for Uganda are in column 8. The pattern is similar to Malawi, though the estimated elasticities are smaller in magnitude. Increases in the labor endowment lead to statistically significant increases in farm labor. The coefficient for ΔE^- is roughly half the magnitude of that for ΔE^+, and is not significant at conventional levels (though the p-value is 0.11). This pattern is necessary for a binding demand constraint, H_t. Unfortunately, we cannot test for differences across activities, because of how the Uganda data were collected.

In Section 5 we discuss these findings, consider some extensions, and further explore the puzzling results for Ethiopia. First, we describe a series of robustness checks for the results of this section.

4.2 Robustness

In this subsection we provide a brief summary of robustness checks for the main results. Details and tables are in Appendix D. The general conclusion is that our core results are robust to various alternative specifications and identification concerns.

In a first set of robustness checks we exclude children under the age of 15 from the definitions of labor demand and the labor endowment. Table S9 recreates Table 3 using this alternative, and the results are broadly consistent with the main results.

Next, we re-estimate our main specifications using changes in owned land as an instrument for changes in cultivated land. As discussed in Section 2.3.2, this is a test for whether non-separation is primarily driven by land market failures. Table S8 shows the results. Somewhat surprisingly, there are no substantive changes in any of the estimated models. While this is not positive evidence that land markets are well functioning in the study countries, it does suggest that non-separation in the panel is not generally due to an inability to adjust the amount of land under cultivation.

In Section 2.4 we described robustness checks in which we use the enumeration area average change in labor endowment (excluding one's own change), $\Delta\overline{E}_{et}^{-h}$, as an instrument for ΔE_{ht}. These IV specifications address the concern that changes in the labor endowment might be endogenous to local labor market conditions. First-stage results, in Appendix B, show that $\Delta\overline{E}_{et}^{-h}$ is an extremely weak instrument, using a range of possible specifications. Hence, the OLS results of the previous subsection are our preferred estimates.

Finally, we noted above that a small number of households in each survey report zero labor demand for some cultivation activities. We cannot assume that all of these entries are measurement error. Findings in Table 3 are based on setting farm labor to 0.1 person-days when it is 0 for one activity but not for others. Our results are broadly consistent if we instead use 0.01 or 1 as a replacement value. The same is true if we use a replacement value of 0.01 for the acreage of the few plots that have zero reported acreage (which is clearly measurement error), instead of dropping those plots entirely. The five additional output tables for these different data cleaning choices are available upon request.

4.3 Heterogeneity by gender and AEZ

In Section 2.5 we described two dimensions of heterogeneity analysis, by gender and agro-ecological zone (AEZ). In Table 4 we report estimates of (11) and (12) that allow for gender-specific labor endowments. Across specifications, coefficient estimates are roughly half the magnitude of those in Table 3. This is because increases (decreases) in the endowment of one gender may be partially offset by decreases (increases) in the endowment of the other (which is not possible when we construct a single measure of ΔE_{ht} for both men and women). The fact that changes in gender-specific endowments (in Table 4) are associated with smaller changes in farm labor utilization than changes in total endowments (in Table 3) indicates that labor markets cannot be entirely segmented by gender. Nevertheless, there are some intriguing patterns of gender differences.

For Ethiopia, symmetric non-separation (column 1) is driven primarily by women entering and leaving the household. Estimated elasticities for $\Delta Female$ are positive and statistically significant; those for men are less than half the magnitude, and not statistically

Table 4: Testing the separation hypothesis separately by gender

Dependent variable:	Δ Log of farm labor (person-days)							
	Ethiopia		Malawi		Tanzania		Uganda	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reference $\times \Delta Male$	0.118		0.215***		0.145		0.088*	
	(0.239)		(0.003)		(0.128)		(0.051)	
Reference $\times \Delta Male^{+}$		0.034		0.243***		0.240**		0.115*
		(0.825)		(0.008)		(0.043)		(0.055)
Reference $\times \Delta Male^{-}$		0.180		0.185*		0.052		0.062
		(0.159)		(0.081)		(0.685)		(0.264)
Reference $\times \Delta Female$	0.258**		0.218*		0.331***		0.042	
	(0.028)		(0.074)		(0.002)		(0.402)	
Reference $\times \Delta Female^{+}$		0.242		0.263*		0.248*		0.022
		(0.142)		(0.070)		(0.089)		(0.726)
Reference $\times \Delta Female^{-}$		0.265*		0.150		0.431***		0.064
		(0.066)		(0.462)		(0.004)		(0.332)
Harvest $\times \Delta Male$	0.078		0.122		0.139			
	(0.413)		(0.108)		(0.111)			
Harvest $\times \Delta Male^{+}$		0.079		0.023		0.261**		
		(0.593)		(0.814)		(0.022)		
Harvest $\times \Delta Male^{-}$		0.080		0.256**		0.023		
		(0.508)		(0.018)		(0.835)		
Harvest $\times \Delta Female$	0.260**		0.099		0.309***			
	(0.029)		(0.403)		(0.004)			
Harvest $\times \Delta Female^{+}$		0.122		0.148		0.237		
		(0.452)		(0.301)		(0.103)		
Harvest $\times \Delta Female^{-}$		0.353**		0.029		0.398***		
		(0.016)		(0.891)		(0.008)		
Weeding $\times \Delta Male$					0.186**			
					(0.033)			
Weeding $\times \Delta Male^{+}$						0.397***		
						(0.000)		
Weeding $\times \Delta Male^{-}$						-0.011		
						(0.927)		
Weeding $\times \Delta Female$					0.395***			
					(0.000)			
Weeding $\times \Delta Female^{+}$						0.280**		
						(0.046)		
Weeding $\times \Delta Female^{-}$						0.528***		
						(0.001)		
Reference-Male: Inc = Dec	.	0.454	.	0.659	.	0.228	.	0.460
Reference-Female: Inc = Dec	.	0.908	.	0.648	.	0.374	.	0.600
Harvest-Male: Inc = Dec	.	0.993	.	0.097	.	0.083	.	.
Harvest-Female: Inc = Dec	.	0.237	.	0.652	.	0.435	.	.
Weeding-Male: Inc = Dec	.	.	.	.	.	0.003	.	.
Weeding-Female: Inc = Dec	.	.	.	.	.	0.241	.	.
Reference: Male=Female	0.421	.	0.989	.	0.267	.	0.578	.
Harvest: Male=Female	0.283	.	0.888	.	0.278	.	.	.
Weeding: Male=Female	.	.	.	.	0.171	.	.	.
Reference: Male=Female,+	.	0.391	.	0.914	.	0.971	.	0.353
Reference: Male=Female,−	.	0.688	.	0.885	.	0.086	.	0.985
Harvest: Male=Female,+	.	0.854	.	0.501	.	0.906	.	.
Harvest: Male=Female,−	.	0.173	.	0.388	.	0.064	.	.
Weeding: Male=Female,+	.	.	.	.	.	0.538	.	.
Weeding: Male=Female,−	.	.	.	.	.	0.009	.	.
Observations	5,650	5,650	4,190	4,190	11,685	11,685	5,362	5,362
R^2	0.028	0.028	0.151	0.152	0.027	0.028	0.089	0.090

Notes: Authors' calculations from LSMS-ISA data. Estimates are marginal effects, interpretable as elasticities. p-values in parentheses. Significance: *** 0.01, ** 0.05, * 0.1. Standard errors clustered by enumeration area. All regressions include year effects and controls for changes in demographic shares. The "Reference" activity is "All non-harvest" for Ethiopia and Malawi, "Planting / land preparation" for Tanzania, and "All farming activities" for Uganda. Labor endowment LE is measured in logs. ΔE^{+} is the change in labor endowment if positive, 0 otherwise; ΔE^{-} is the change in labor endowment if negative, 0 otherwise. The dependent variable and LE are defined so that everyone over age 15 is 1 worker, and children aged 11-14 count as $(Age - 10) \times 0.2$ workers.

different from zero. This pattern is maintained in column 2, which allows for asymmetry. Decreases in endowments remain the drivers of non-separation in Ethiopia, yet the loss of a female has a much greater impact on farm labor than the loss of a male. The difference in coefficient magnitudes is particularly substantial in the harvest period, when it approaches statistical significance (p-value = 0.17). This is suggestive not only of partially segmented labor markets, but also of a much tighter market for female labor.

Gender differences are less pronounced for Malawi. Results are in columns 3 and 4 of Table 4. The only surprising estimate is for "Harvest $\times\Delta Male^{-}$", which is positive and statistically significant. This is consistent with a binding supply constraint for male labor during the harvest, whereas in Table 3 the overall pattern for Malawi is one of a binding demand constraint.

The gender-differentiated estimates for Tanzania show the most striking pattern. In the symmetric specifications (column 5), changes in the endowment of women are associated with much larger changes in farm labor than are changes in the endowment of men, in every cultivation phase. When we allow for asymmetry, a second dimension of gender difference emerges: in every cultivation phase, the strongest drivers of changes in farm labor are *increases* in available male labor and *decreases* in available female labor. In all three cultivation phases the relationship between $\Delta Male^{-}$ and farm labor utilization is a tightly estimated zero (recall that those are p-values in parentheses), while that for $\Delta Female^{-}$ is almost equal in magnitude to the aggregate coefficient in column 6 of Table 3. Farm labor does not decrease when a man leaves the household. Following the predictions of Section 2, this is consistent with a general pattern of binding supply constraints for female labor and binding demand constraints for male labor.

Uganda is the only country in which farm labor is more responsive to changes in the endowment of men than the endowment of women. In both columns 7 and 8, coefficients on the change in female labor endowment are near zero, and estimated tightly. The implication is that both symmetric and asymmetric non-separation in Table 3 is driven primarily by men entering and leaving households.

To summarize, across study countries we find a general indication of partially segmented labor markets. There is clearly some substitutability between male and female labor, because estimated elasticities are smaller than those in Table 3, and most of the male-female differences are not statistically significant at conventional levels (p-values from F-tests reported in the lower part of Table 4). The most notable difference in Tanzania, where we see that evidence in Table 3 of a possible labor supply constraint is concentrated in the market for female labor.

Analysis of heterogeneity by agro-ecological zone (AEZ) shows some substantial differences across areas of the study countries. This is in contrast to Dillon and Barrett (2017), who find little evidence of differences by AEZ in cross-sectional tests for symmetric non-separation. To economize on space we provide details of the AEZ analysis in Appendix section D.3. The general takeaway is that the nationally representative averages in Table 3 are useful for understanding the prevailing labor market conditions at the country level, but allowing for differences by agro-climate can sharpen our understanding of the conditions under which non-separation is most egregious.

5 Discussion and extensions

In this final section we interpret our findings, place them in the context of the recent literature, and report a number of extensions to the main analysis.

5.1 Interpretation of results for Malawi, Tanzania, and Uganda

For two of the study countries, Malawi and Uganda, the results satisfy the necessary conditions for a binding labor demand constraint. For Tanzania, the pattern of asymmetries leans toward the same conclusion, although it is weaker. We argued in Section 2.3.1 that in panel data, this pattern of asymmetric responses, with greater responses to ΔE^{+} than ΔE^{-}, is not consistent with non-separation due to credit constraints. Furthermore, in Section 4.2 we reported that the findings are unchanged if we instrument for changes in cultivated acreage using changes in owned acreage. Hence, having ruled out the primary competing interpretations of this asymmetric pattern, we interpret these findings as evidence of binding labor demand constraints facing the average household in Malawi and Uganda, and likely facing many households in Tanzania. In Tanzania, the demand constraint is tighter for men than for women.

This finding does not suggest that there are no other market shortcomings in these countries. It is almost certainly the case that some households face labor supply problems, and yet others enjoy separation. Our finding of a prevailing labor demand constraint is an average across the farming areas of these three countries. Furthermore, our positive finding of a binding labor demand constraint does not preclude other possible market failures. We have ruled out leading roles for credit constraints and missing land markets as the drivers of *asymmetric* non-separation in Malawi, Tanzania, and Uganda, but incomplete markets for these or other inputs, for outputs, or for insurance could all drive symmetric non-separation, and thereby influence the level values of the estimated elasticities in Table 3. We think these results are an important advance because they allow us to link a finding of non-separation to a specific inference about the labor market, but it would be an overstatement to conclude that other markets are necessarily working well.

Widespread evidence of labor demand constraints at the household level suggests that in the aggregate, many local labor markets are characterized by excess supply. This finding is consistent with recent evidence of downwardly sticky nominal wages in other developing country settings (Kaur, 2016), and with evidence of a persistent rural-urban gap in labor productivity in sub-Saharan Africa (Gollin, Lagakos and Waugh, 2014; McCullough, 2017).[9] There is a clear structural interpretation of this finding. In the absence of widespread mechanization, the nature of agriculture production forces all but the most efficient labor markets to exhibit some degree of over-supply at certain times of year. If the market is able to meet demand during the peak periods, the workers who supply that labor are unlikely to be fully utilized during other periods.[10] In this sense, because the shortcoming is on the labor demand side, it is not surprising that the degree of non-separation is greatest during the weeding period in Tanzania and the pre-harvest period in Malawi. Rather than causing the market to "seize up," the additional demand for workers during planting and harvest makes for more efficient use of the available labor supply. During other cultivation phases, when

[9]There is some disagreement in the current literature about whether the rural-urban gap reflects real differences, or unobserved individual productivity. See Hicks et al. (2017).

[10]One common use of under-utilized labor in off-seasons is non-farm enterprises. Recent work by Brummund and Merfeld (2017) analyze how efficiently households allocate labor between farm and non-farm activities, and do not find any evidence of misallocation for Malawian farmers.

there is less demand, farm labor is especially responsive to the household labor endowment, because household members have little else to do.[11]

This raises the question of why we also see non-separation during the peak demand periods in Malawi and Tanzania. There are at least two possibilities. First, as we just noted, other market shortcomings can cause non-separation during any cultivation phase. Second, risk aversion related to the capacity to hire sufficient workers during peak periods might lead households to be cautious about both how much acreage they plant and how intensely they plan to cultivate that acreage (Fafchamps, 1993; Kochar, 1999). This would lead to "over-staffing," on average, through endogenous adjustment of farm size and farming intensity rather than endogenous adjustment of the labor endowment.

Finally, it is noteworthy that coefficient estimates for Uganda are much smaller in magnitude than those for the other countries. One possible explanation is that the agro-ecological conditions in Uganda allow many households to complete two crop cycles per year, which supports the development of more robust local labor markets. A second potential explanation is that in many respects Uganda is the most market-friendly of the study countries, allowing for a generally more efficient allocation of resources (including labor).[12]

5.2 Interpretation and extensions for Ethiopia

The most surprising results in Section 4.1 are for Ethiopia. There, we found greater average responses to ΔE^- than ΔE^+, which is consistent with a binding labor supply constraint. However, as noted in Section 2.3.1, this pattern of asymmetries is also consistent with a binding credit constraint. To distinguish these two hypotheses we re-estimate (11) and (12), interacting all ΔE and $\mathbf{F}$ terms with a binary variable for whether the household is above median wealth (which proxies for access to credit). We do this twice, using two different wealth measures: expenditure-per-capita, which is potentially endogenous to farming choices

[11] This interpretation of the level differences in the pre-harvest and harvest elasticities for Malawi may seem at odds with the planting-season spike in labor utilization shown in Figure 2. However, the pre-harvest category in Malawi is an average result between a brief planting period of intense demand, and a long cultivation period of much greater slack. It seems to be the case that the latter has a greater effect on the estimated "Reference $\times\Delta E^+$" coefficient.

[12] For instance, Uganda is the highest-ranked of the study countries in the 2017 World Bank Doing Business Index (though none of the countries performs very well).

but which closely tracks liquidity; and an asset index based on ownership of durables and household dwelling characteristics, which is likely exogenous to current year agricultural outcomes but is a noisier measure of credit access.

A third possible interpretation of the asymmetry in Ethiopia is that the Productive Safety Net Programme (PSNP), a major workfare program operating during the study period, crowded out supply of labor to the private market. The PSNP targets 7-8 million people, in three stages.[13] The goal is to provide timely and adequate support to food insecure households by generating remunerative work during the lean season, which is typically when cultivation of the next crop is under way. Clearly, PSNP crowd-out of supply to the private market would be an unintended negative consequence of the program. To test the hypothesis that the PSNP is crowding out the supply of workers to farms, we use a similar set of interactions to allow for heterogeneity across areas where the PSNP is and is not active at the woreda level.[14]

Table 5 shows the results. Columns 1 and 2 allow for heterogeneity by PSNP activity. If supply of labor to PSNP was causing the asymmetry found in the main analysis, we would see stronger evidence of non-separation in the woredas with PSNP (bottom panel) than in woredas with no PSNP activity. The results in Table 5 do not show that pattern. There is evidence of non-separation in both PSNP and no-PSNP woredas, and in both areas the association between ΔE^- and changes in farm labor is stronger than that for ΔE^+. Hence, the PSNP does not seem to be causing the binding labor supply constraint for Ethiopia.

Results allowing for heterogeneity by wealth are in Columns 3-6 of Table 5. In columns 3 and 4, which split the sample into above/below median expenditure-per-capita, we see an extremely asymmetric response for the poor households (top panel), consistent with either a binding labor supply or binding credit constraint. In constrast, non-separation for the above median wealth households is symmetric (lower panel). The general pattern is the same in columns 5 and 6, which split the sample based on assets rather than expenditure.

[13]First, food-insecure woreda (districts) are identified as those located in rural regions that have received food aid for a significant period. Funds are distributed in proportion to the number of food-insecure people in each woreda. Second, a local Woreda Council allocates transfers to each village (kebele) in that woreda. Third, a Kebele Council allocates the transfers to individuals in the village.

[14]The LSMS-ISA survey contains details about PSNP participation, which allow us to distinguish between participating and non-participating areas.

Table 5: Testing possible causes of asymmetric non-separation in Ethiopia

Dependent variable:	Δ Log of farm labor (person-days)					
	by PSNP		by Expenditure		by Assets	
	(1)	(2)	(3)	(4)	(5)	(6)
	No PSNP		**Wealth** $< 50\%$		**Wealth** $< 50\%$	
Reference $\times \Delta E$	0.389*		0.280		0.258	
	(0.052)		(0.226)		(0.286)	
Reference $\times \Delta E^+$		0.243		-0.531		-0.498
		(0.464)		(0.123)		(0.170)
Reference $\times \Delta E^-$		0.495*		0.915**		0.879**
		(0.094)		(0.017)		(0.014)
Harvest $\times \Delta E$	0.407**		0.256		0.131	
	(0.045)		(0.241)		(0.577)	
Harvest $\times \Delta E^+$		0.128		-0.450		-0.181
		(0.707)		(0.168)		(0.657)
Harvest $\times \Delta E^-$		0.616*		0.808**		0.377
		(0.080)		(0.030)		(0.277)
Reference: Inc = Dec	.	0.600	.	0.012	.	0.011
Harvest: Inc = Dec	.	0.386	.	0.024	.	0.340
Reference = Harvest	0.885	.	0.857	.	0.429	.
Reference = Harvest, +	.	0.559	.	0.703	.	0.257
Reference = Harvest, −	.	0.580	.	0.609	.	0.026
	PSNP		**Wealth** $\geq 50\%$		**Wealth** $\geq 50\%$	
Reference $\times \Delta E$	0.603**		0.624***		0.713***	
	(0.031)		(0.002)		(0.001)	
Reference $\times \Delta E^+$		0.152		0.657**		0.990***
		(0.677)		(0.048)		(0.005)
Reference $\times \Delta E^-$		0.974**		0.588**		0.494
		(0.030)		(0.035)		(0.132)
Harvest $\times \Delta E$	0.506*		0.594***		0.794***	
	(0.072)		(0.008)		(0.000)	
Harvest $\times \Delta E^+$		0.311		0.597		0.610*
		(0.445)		(0.100)		(0.081)
Harvest $\times \Delta E^-$		0.664		0.581*		0.918**
		(0.145)		(0.088)		(0.011)
Reference: Inc = Dec	.	0.186	.	0.880	.	0.360
Harvest: Inc = Dec	.	0.596	.	0.977	.	0.582
Reference = Harvest	0.585	.	0.840	.	0.541	.
Reference = Harvest, +	.	0.627	.	0.830	.	0.032
Reference = Harvest, −	.	0.202	.	0.978	.	0.030
Reference: Top = Bottom	0.528	.	0.204	.	0.138	.
Reference: Top = Bottom, +	.	0.856	.	0.012	.	0.004
Reference: Top = Bottom, −	.	0.368	.	0.423	.	0.397
Harvest: Top = Bottom	0.773	.	0.242	.	0.033	.
Harvest: Top = Bottom, +	.	0.733	.	0.031	.	0.144
Harvest: Top = Bottom, −	.	0.932	.	0.611	.	0.221
Observations	5,650	5,650	5,650	5,650	5,650	5,650
R^2	0.031	0.032	0.032	0.035	0.032	0.033

Notes: Authors' calculations from LSMS-ISA data. Estimates are marginal effects, interpretable as elasticities. p-values in parentheses. Significance: *** 0.01, ** 0.05, * 0.1. Standard errors clustered by enumeration area. All regressions include year effects and controls for changes in demographic shares. The "Reference" activity is "All non-harvest." Labor endowment LE is measured in logs. ΔE^+ is the change in labor endowment if positive, 0 otherwise; ΔE^- is the change in labor endowment if negative, 0 otherwise. The dependent variable and LE are defined so that everyone over age 15 is 1 worker, and children aged 11-14 count as $(Age - 10) \times 0.2$ workers.

If asymmetric non-separation in Ethiopia was driven by a physical limitation on the pool of available workers at the prevailing wage, we would expect all households to be affected. Instead, we see a sharp divide between households along wealth lines. The strong suggestion is that a lack of credit, rather than some other feature specific to the labor market, is responsible for the binding labor supply constraint for poor households in Ethiopia.

5.3 Concluding comments

The separation test of Benjamin (1992) is a well-established method for testing the completeness of rural markets. LaFave and Thomas (2016) recently placed this classic test back on the research frontier, by showing that Benjamin's finding of separation disappears with more comprehensive data and controls for household fixed effects. The main contribution of those papers is to inform researchers of the appropriate framework for modeling rural households. Lafave and Thomas cite numerous papers that assume separation in order to justify an exclusive focus on either the profit-maximization problem of the household business or the utility maximization problem of the household members. They suggest that such assumptions may not be justifiable.

Like LaFave and Thomas (2016), we find evidence of non-separation in both pooled and fixed effects regressions. We then advance the literature in two ways. First, we theoretically develop a test for asymmetric non-separation, and derive conditions under which asymmetric responses to increases and decreases in the labor endowment can reveal something about the underlying labor market conditions. Second, we implement our tests in nationally representative data from four countries. We show that some important differences across countries do not emerge until one allows for asymmetric non-separation at different times of year, and that additional nuances emerge when we allow for possible differences by gender and by agro-ecological zone. This underscores the risks of taking a result from one country, or one region of one country, and treating it as representative for all of sub-Saharan Africa.

References

Ali, Daniel Ayalew, Klaus Deininger, and Markus Goldstein. 2014. "Environmental and gender impacts of land tenure regularization in Africa: Pilot evidence from Rwanda." *Journal of Development Economics*, 110: 262–275.

Allendorf, Keera. 2007. "Do Women's Land Rights Promote Empowerment and Child Health in Nepal?" *World Development*, 35(11): 1975–1988.

Barrett, Christopher B. 1996. "On price risk and the inverse farm size-productivity relationship." *Journal of Development Economics*, 51(2): 193–215.

Benjamin, Dwayne. 1992. "Household composition, labor markets, and labor demand: testing for separation in agricultural household models." *Econometrica*, 60(2): 287–322.

Besley, Timothy. 1995. "Property Rights and Investment Incentives: Theory and Evidence from Ghana." *Journal of Political Economy*, 103(5): 903.

Brummund, Peter, and Joshua D Merfeld. 2017. "Allocative Efficiency of Non-Farm Enterprises in Agricultural Households: Evidence from Malawi." *mimeo.*

Dillon, Brian, and Alessandra Voena. 2016. "Inheritance Customs and Agricultural Investment." University of Washington working paper.

Dillon, Brian, and Christopher B. Barrett. 2017. "Agricultural factor markets in Sub-Saharan Africa: an updated view with formal tests for market failure." *Food Policy*, 67: 64–77.

Doss, Cheryl, Chiara Kovarik, Amber Peterman, Agnes Quisumbing, and Mara van den Bold. 2015. "Gender inequalities in ownership and control of land in Africa: myth and reality." *Agricultural Economics*, 46(3): 403–434.

Dreze, Jean P, and Anindita Mukherjee. 1989. "Labour contracts in rural India: Theories and evidence." In *The Balance between Industry and Agriculture in Economic Development.* Vol. 3, , ed. Sukhamoy Chakravarty, 233–265. Macmillan Press Ltd.

Fafchamps, Marcel. 1993. “Sequential labor decisions under uncertainty: An estimable household model of West-African farmers.” *Econometrica*, 61(5): 1173–1197.

Feder, Gershon. 1985. “The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints.” *Journal of Development Economics*, 18(2): 297–313.

Goldstein, Markus, and Christopher Udry. 2008. “The profits of power: Land rights and agricultural investment in Ghana.” *Journal of Political Economy*, 116(6): 981–1022.

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

Harris, John R, and Michael P Todaro. 1970. “Migration, unemployment and development: a two-sector analysis.” *The American Economic Review*, 60(1): 126–142.

Hicks, Joan Hamory, Marieke Kleemans, Nicholas Y Li, and Edward Miguel. 2017. “Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata.” *NBER Working Paper*, 23253.

Kaur, Supreet. 2016. “Nominal wage rigidity in village labor markets.” *NBER Working Paper*, 20770.

Kochar, Anjini. 1999. “Smoothing consumption by smoothing income: hours-of-work responses to idiosyncratic agricultural shocks in rural India.” *Review of Economics and Statistics*, 81(1): 50–61.

Kumar, Neha, and Agnes R. Quisumbing. 2012. “Gendered Impacts of the 2007-08 Food Price Crisis: Evidence Using Panel Data from Rural Ethiopia.” *Food Policy*, 38: 11–22.

LaFave, Daniel, and Duncan Thomas. 2016. “Farms, families, and markets: New evidence on agricultural labor markets.” *Econometrica*, 84(5): 1917–60.

Le, Kien T. 2010. “Separation hypothesis tests in the agricultural Household model.” *American Journal of Agricultural Economics*, aaq070.

Lewis, W Arthur. 1954. "Economic development with unlimited supplies of labour." *The Manchester School of Economic and Social Studies*, 22(2): 139–191.

McCullough, Ellen B. 2017. "Labor productivity and employment gaps in Sub-Saharan Africa." *Food Policy*, 67: 133–152.

Osmani, Siddiqur Rahman. 1990. "Wage determination in rural labour markets: the theory of implicit co-operation." *Journal of Development Economics*, 34(1): 3–23.

Sen, Amartya K. 1966. "Peasants and Dualism with or without Surplus Labor." *Journal of Political Economy*, 74(5): 425–450.

Singh, Inderjit, Lyn Squire, John Strauss, et al. 1986. *Agricultural household models: extensions, applications, and policy.* Johns Hopkins University Press.

Ubink, Janine M, and Julian F Quan. 2008. "How to combine tradition and modernity? Regulating customary land management in Ghana." *Land Use Policy*, 25(2): 198–213.

Udry, Christopher. 1996. "Gender, Agricultural Production, and the Theory of the Household." *Journal of Political Economy*, 104(5): 1010.

Udry, Christopher. 1999. "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture." In *Trade, Growth and Development: Essays in Honor of T.N. Srinivasan.* , ed. Ranis and Raut. Amsterdam:Elsevier Science.

Yngstrom, Ingrid. 2002. "Women, Wives and Land Rights in Africa: Situating Gender Beyond the Household in the Debate Over Land Policy and Changing Tenure Systems." *Oxford Development Studies*, 30(1): 21–40.

A Data Appendix

A.1 Description of LSMS-ISA surveys

The Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) is a household survey project designed to improve the quality and consistency of agricultural data in sub-Saharan Africa. Enabled by funding from the Bill and Melinda Gates Foundation, the World Bank LSMS team partnered with national statistics offices to design and implement the nationally representative household surveys which had a strong focus on agriculture. The surveys had a multi-topic approach to allow the analysis of links between agriculture, socioeconomic status, and non-farm income activities. Our paper used the LSMS-ISA data for countries which had panel data available at the start of the project, Ethiopia, Malawi, Tanzania, and Uganda.

For Ethiopia, the LSMS-ISA project supported the Ethiopia Socioeconomic Survey (ERSS), which was implemented by the Central Statistical Agency of Ethiopia (CSA). Wave 1 of the ERSS was conducted from September 2011 to March 2012 and comprised 4,000 households in rural and small towns across Ethiopia. The sample was constructed in two stages, with each enumeration area (EA) being selected based probability proportional to size of population (PPS), and then within each enumeration area, households were selected using simple random sampling. For Wave 1, 290 rural EAs were selected and 43 small town EAs were selected. Wave 2 for Ethiopia was conducted between September 2013 and April 2014.

In Malawi, the LSMS-ISA project partnered with the Malawi Integrated Household Survey (IHS) Program which is administered by the National Statistics Office (NSO), starting with the third wave (IHS3). This third wave of the IHS was conducted from March 2010 to March 2011, and comprised 12.271 households. The overall sample is representative at the district, regional, urban-rural, and national levels. A subset of the original sample was designated to be panel households prior to the start of the IHS3 fieldwork. These 3,247 households were re-interviewed as part of the Integrated Household Panel Survey in 2013.

In Tanzania, the LSMS team partnered with the National Bureau of Statistics (NBS) to establish the Tanzania National Panel Survey (TZNPS). The first wave of TZNPS was conducted from October 2008 to September 2009 and comprised 3,265 households. All of these households were targeted to be included in Waves 2 and 3. Wave 2 was conducted between October 2010 and December 2011, and wave 3 was conducted between October 2012 and December 2013.

For Uganda, the LSMS team supported the Uganda National Panel Survey (UNPS), which is administered by the Uganda Bureau of Statistics (UBoS). The UNPS built on the Uganda National Household Survey (UNHS), which was conducted in 2005 and 2006. The first round of the LSMS supported UNPS was conducted in 2009 and 2010 and comprised 3,200 households. All of these households were apart of the UNHS, and are representative at the national, urban/rural, and main regional levels. The UNPS was collected again in the following two years (2009/10 and 2010/11) to comprise four rounds of data.

A.2 Construction of variables

Here we put details about the construction of the labor endowment and labor demand (utilization) variable for each country. The labor endowment variable measures the number people in each household that could possibly participate in agricultural activities. The labor demand variable measures the number of days that were worked on each household's agricultural activities. This includes some combination of household labor, hired labor, and free/exchange labor, depending on the country. The labor demand was also captured for separate agricultural activities in some of the countries.

Agricultural data for Ethiopia was collected through two separate surveys, one post-planting and one post-harvest. The post-planting questionnaire asks about all non-harvest cultivation activities, such as land preparation, planting, ridging, weeding, and fertilizing, while the post-harvest survey asks about labor used for harvesting and threshing. From these two separate questions we form separate variables for farm labor demand during "Cultivation" and "Harvest." Data for the other study countries were collected in a single survey. For Malawi, households could report their own labor separately for land preparation/planting, weeding/fertilizing, and harvesting. But reporting of hired labor was divided between "all non-harvest" and "harvest" activities. Hence, to form the variables for total quantity of labor demanded on farms we combine the first two categories of household labor with the "all non-harvest" hired labor to form a variable for "Cultivation", and sum the household and hired labor to form the "Harvest" variable.

Some areas of Tanzania have two rainy seasons. We only use data from the primary farming period, the long rainy season, which runs from January/February to April/May. All three waves of the Tanzania data record household and hired labor for preparation, weeding, and harvest activities. Waves 2 and 3 also record fertilizing labor, but we exclude that as it appears to not be included in any of the Wave 1 categories. The Wave 1 data for Uganda records labor separately by agricultural activity, but Waves 2-4 only record total labor, so we combine all the Wave 1 activities together for consistency.

The countries also differ in how they capture free/exchange labor and labor provided by children. For Ethiopia, the data identifies children who worked on the farm as being from the household, hired, or working as free/exchange labor. The Ethiopia data also records the ages of each child from the household that worked on the farm, so their labor is phased-in to total labor demand in the same as they are into the household labor endowment, with work provided by children aged 10 and under not being counted, each day worked by an 11 year old counted at 20%, 40% for 12 year olds, and so on until the work of 15 year olds is counted fully. The ages of children that worked on the farm, but were not members of the household, were not recorded. All of the non-household child labor was discounted 50%.

Like Tanzania, some areas of Malawi have two rainy seasons. We only use data for the long rainy season, which is the main cultivation period. The Malawi data recorded children in a similar fashion as Ethiopia. The ages of children from the household were recorded, so their labor demand was phased-in, whereas the ages of children from outside the household were not recorded and their labor was discounted by 50%. The Malawi data also does not separately record outside workers who may have worked as free/exchange laborers.

For household labor in both Ethiopia and Malawi, the survey records the number of weeks worked by each individual in the reference period, the average number of days worked in each week, and the average number of hours worked each day. We construct the total hours worked by that individual on that plot by multiplying those three values together. However, Tanzania and Uganda both record household labor in days. To standardize the units of labor demand across datasets, we divide total hours worked by 6 to obtain the total days worked for household workers in Ethiopia and Malawi.

For the Tanzania data, children hired from outside the household are only recorded separately in Waves 2 and 3. Ages are recorded for children from within the household who work on the farm, and so their time is phased in. Ages for children hired from outside the household are not recorded, and so their time is discounted by 50%. The Tanzania data does not have a separate variable for free/exchange labor, but the Wave 2 and 3 surveys include specific guidance to include labor paid with in-kind payments.

For the Uganda data, children working on the farm from the household are only captured in Wave 1, whereas children hired from outside the household are recorded in all four waves. Wave 4 for Uganda also separately records free/exchange labor, which we include in the total labor demand as it appears to have been counted as hired labor in the previous waves.

Another issue with the Uganda data is that the survey only provides details for up to three workers on each plot. The total amount of days worked on each plot is recorded, but if more than three family members worked on a plot, we are not able to identify who they are. This means that we are not able to separate out children from adults for household workers, and need to treat all household labor equally.

A.3 Details for Figure 2

In this subsection we provide details for construction of the graph in panel A of Figure 2. All variables used to construct this figure were reported at the household-plot level. Respondents reported the month in which planting ended, and whether planting ended in the first half or second half of the month. To estimate the beginning of the planting period, we took the maximum number of weeks that any household member worked for land preparation or planting on the plot, and rounded to the nearest half-month, letting 4 weeks equal a month. If household planting labor on the plot was reported to be zero (7% of plots), we assigned the plot the shortest possible planting duration, 0.5 months. Households reported the month in which harvesting began, and the month in which harvesting ended. We estimated the harvest duration as 0.5*(difference+1), where difference = end month - start month. This set the harvest period duration equal to its expected value, given information on the start and end month only. We then randomly assigned plots to begin the harvest at the beginning or middle of the reported start month. Based on these estimates, a small number of plots (<3%) reported an end of planting period that occurred after the beginning of the harvest period. While this is possible on multi-cropped plots, we dropped these plots because of suspected misreporting. We then assigned the end of the planting period to be the beginning of the cultivation (weeding and fertilizing) period, and the beginning of the harvest period to be the end of the cultivation period. Household labor supply to the plot was reported separately for planting, cultivation, and harvest. We evenly divided the total person-days for each period between the half-months in that period. Hired labor supply to the plot was reported as an aggregate figure for planting and cultivation, with harvest reported separately. We evenly divided the planting and cultivation labor between the half-months in those two periods, and assigned harvest labor to that period in the same manner as the household harvest labor. Note that if hired labor follows a pattern similar to household labor, with more intensive application at planting, then this approach will underestimate planting labor and overestimate cultivation labor. Child labor is excluded from these estimates as it was not collected for the hired workers, but this is a negligible fraction of total labor. We assume that a full work day is 6 hours long. The figure shown is a local polynomial regression of labor demand (in person-days per half-month) on time, using an Epanechnikov kernel.

B Instrumenting for changes in labor endowments

In this section we describe a sequence of empirical models that use the village level average change in labor endowment (excluding the household h change), $\Delta \overline{E}_{et}^{-h}$, as an instrument for ΔE_{ht}. The idea underlying this approach is that if households are endogenously adjusting their labor endowments in response to local labor market conditions, we should see correlated changes in ΔE_{ht} for households facing a common local labor market. Because each household has only a single measure of ΔE_{ht}, we use a second stage specification with symmetric response pooled across cultivation activities, like equation (10). In practice, the second stage does not matter, because the first stage reveals that the proposed instrument is very weak.

The first stage equation is as follows:

$$\Delta E_{ht} = \beta_0 + \beta_1 \Delta \overline{E}_{et}^{-h} + \gamma A_{ht} + \gamma_2 demog_{ht} + \nu_t + \epsilon_{fht} \tag{14}$$

We estimate (14) in a 2SLS framework, with the second-stage represented by equation (10). For the first set of estimates we exclude children aged 11-14 from the definition of the labor endowment, imposing a binary cut-off at age 15. This reflects the nature of the instrument, which is based on other households' net positions in the local market. Next, we use the equivalence scale described in Section 3, with children aged 11-14 represented as $(Age - 10) * 0.2$ working age equivalents in the definition of the labor endowment. We include these children in ΔE_{ht} but not $\Delta \overline{E}_{et}^{-h}$, to represent the idea that a household might take into account its own aging children when making endogenous adjustments to labor. Finally, because a handful of enumeration areas have only a few households, in a third set of estimates we drop any areas with fewer than 6 households to ensure that $\Delta \overline{E}_{et}^{-h}$ is not susceptible to small sample biases.

Table S1: F-statistics from first-stage IV regressions using $\Delta \overline{E}_{et}^{-h}$ as instrument

	Excluding children (1)	Including children (2)	Only EAs with 6+ households (3)
Ethiopia	3.49	6.54	4.77
Malawi	3.27	3.77	3.59
Tanzania	0.64	0.04	1.10
Uganda	1.35	2.99	0.70

Notes: Authors' calculations from LSMS-ISA data.

Table S1 shows the first stage F-statistics for the instrument, for each of the three estimation approaches. In the main analysis, column 1, the largest F-stat is 3.49. Three of the F-stats are larger in column 2, but the largest is still only 6.54. Limiting the analysis to enumeration areas with at least 6 households, in column 3, does not lead to improvements in the power of the instrument.

The clear takeaway from this analysis is that households in the same geographical area do not exhibit correlated changes in labor endowments. We treat this as evidence that changes in household composition are largely exogenous to local labor market conditions. Another way to see this is to examine the spatial distribution of net changes in the household labor endowment.

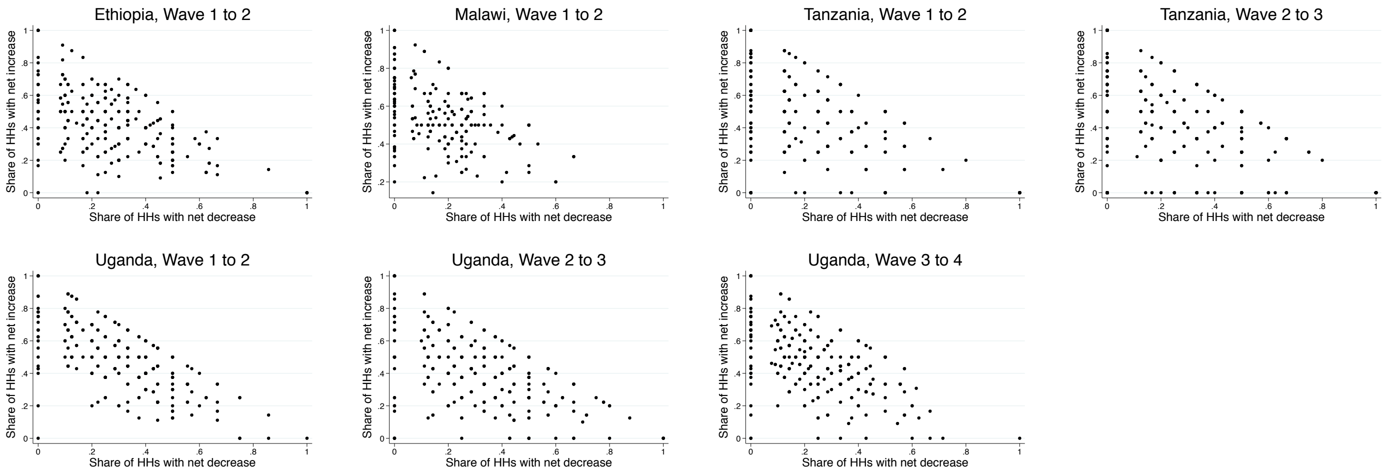


Figure S1: Distribution of net labor endowment changes, village level

Notes: Authors' calculations from LSMS-ISA data. Each dot represents a village or enumeration area. Sample size per village is roughly equivalent in most surveys, so these patterns are similar if we use circles to represent the size of each village sample.

Figure S1 shows location-level scatter plots of the proportion of households with a net increase (vertical axis) and a net decrease (horizontal axis) in the labor endowment, by survey wave. The points are defined at the enumeration area level for all countries. By construction, all points lie in the unit simplex. The clear pattern in all surveys is of clustering in the interior of the simplex, rather than along the axes. When there is clustering, it is along the vertical axis, and is usually driven by the aging of children aged 11-14. The overall indication is that most enumeration areas contain a mix of some households that experience a net increase in labor endowment, and some that experience a net decrease. The general pattern in these figures does not change if we represent each enumeration area with a bubble scaled to represent the number of households represented.

C Descriptive statistics for migrants and stayers

In this section we examine the labor supply to the farm by new household members, incumbents (stayers), and those who move out. We refer to non-stayers as migrants. We also look at the summary statistics for migrants and stayers in order to characterize the population responsible for changes in labor endowments.

Table S2: Summary statistics for migrants and stayers: Ethiopia

	(1) Moved-Out mean/sd	(2) Moved-In mean/sd	(3) Stayed mean/sd
Age	26.25	29.05	22.28
	(15.36)	(16.73)	(18.44)
% Male	0.51	0.45	0.51
	(0.50)	(0.50)	(0.50)
% Child	0.60	0.34	0.57
	(0.49)	(0.47)	(0.50)
% Other Family Member	0.35	0.57	0.43
	(0.48)	(0.50)	(0.49)
% Not a Family Member	0.06	0.07	0.00
	(0.23)	(0.26)	(0.04)
% Worked Cultivation	0.54	0.35	0.42
	(0.50)	(0.48)	(0.45)
Avg. Days Worked Cultivation	66.11	43.32	59.81
	(117.96)	(52.97)	(98.43)
% Worked Harvest	0.55	0.50	0.46
	(0.50)	(0.50)	(0.44)
Avg. Days Worked Harvest	23.10	27.22	30.00
	(29.32)	(37.02)	(32.84)
Number of Obs.	1,084	544	13,180

Notes: Authors' calculations from LSMS-ISA data.

Table S2 shows migrant and stayer characteristics for Ethiopia. Columns 1 and 2 differentiate between out-migrants and in-migrants. The table uses data from two waves. Column 1 uses data from Wave 1, before the individual moved-out, and column 2 only uses data from Wave 2, after the migrant moved-in. Descriptive statistics for stayers, in column 3, are averages across Waves 1 and 2. Essentially none of the between-column differences are statistically significant. There are two notable takeaways. First, the large majority of both in- and out-migrants are family members. The average migrant is in his or her late 20s, with balance across genders. It is likely that many of these changes in the labor endowment are due to marriage. Second, in-migrants and out-migrants do not do substantially more or less work on farm than stayers. Rather, they seem to fit into the working life of the household when they are a part of it, working on the family farm much like everyone else.

In Table S3 we show the same statistics for Malawi. The characteristics of migrants and stayers are remarkably similar to those for Ethiopia. The only noteworthy difference between the countries is that an even smaller percentage of migration in Malawi is by non-

Table S3: Summary statistics for migrants and stayers: Malawi

	(1) Moved-Out mean/sd	(2) Moved-In mean/sd	(3) Stayed mean/sd
Age	28.36	29.26	22.29
	(16.05)	(15.23)	(18.61)
% Male	0.50	0.51	0.49
	(0.50)	(0.50)	(0.50)
% Child	0.59	0.31	0.58
	(0.49)	(0.46)	(0.49)
% Other Family Member	0.40	0.65	0.41
	(0.49)	(0.48)	(0.49)
% Not a Family Member	0.01	0.02	0.00
	(0.09)	(0.12)	(0.03)
% Worked Cultivation	0.55	0.48	0.45
	(0.50)	(0.50)	(0.45)
Avg. Days Worked Cultivation	19.06	16.80	19.49
	(19.78)	(15.34)	(14.59)
% Worked Harvest	0.41	0.54	0.45
	(0.49)	(0.50)	(0.42)
Avg. Days Worked Harvest	9.83	7.44	14.63
	(16.10)	(10.79)	(12.05)
Number of Obs.	836	701	9,992

Notes: Authors' calculations from LSMS-ISA data.

family members. Once again the differences between the three columns are not statistically significant.

In Table S4 we show the same set of summary statistics, for Tanzania. The first three columns display the results for changes between Waves 1 &2, and the last three columns for changes between Waves 2 & 3. Once again, the third and sixth columns show the characteristics of stayers pooled across both waves. For Tanzania the most important patterns from Ethiopia and Malawi are maintained. It is notable, and somewhat surprising, that more women than men migrate in Tanzania. This could reflect between-country variation in customs governing the residency location of couples after marriage. However, the finding that both in- and out-migrants work on the farm to a similar extent as everyone else is as true in Tanzania as in Ethiopia and Malawi.

Finally, in Tables S5, S6, and S7, we show the similar statistics for Uganda. However, in Uganda, the data does identify the specific person who worked on each plot in Wave 1, so column 1 of Table S5 can not show what percent over out-migrants worked on the farm.

Table S4: Summary statistics for migrants and stayers: Tanzania

	Wave 1-2			Wave 2-3		
	(1)	(2)	(3)	(4)	(5)	(6)
	Moved-Out	Moved-In	Stayed	Moved-Out	Moved-In	Stayed
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
Age	28.44	28.71	24.02	29.60	29.77	24.32
	(17.10)	(15.74)	(20.05)	(17.72)	(16.68)	(20.34)
% Male	0.37	0.38	0.49	0.41	0.32	0.50
	(0.48)	(0.48)	(0.50)	(0.49)	(0.47)	(0.50)
% Child	0.44	0.28	0.48	0.40	0.27	0.49
	(0.50)	(0.45)	(0.50)	(0.49)	(0.45)	(0.50)
% Other Family Member	0.51	0.65	0.52	0.55	0.62	0.51
	(0.50)	(0.48)	(0.50)	(0.50)	(0.48)	(0.50)
% Not a Family Member	0.04	0.07	0.00	0.05	0.10	0.01
	(0.21)	(0.25)	(0.05)	(0.21)	(0.30)	(0.08)
% Worked Cultivation	0.49	0.39	0.44	0.47	0.41	0.44
	(0.50)	(0.49)	(0.44)	(0.50)	(0.49)	(0.44)
Avg. Days Worked Cultivation	27.18	23.38	28.10	22.79	23.70	24.07
	(35.87)	(19.68)	(25.82)	(20.06)	(21.91)	(17.79)
% Worked Harvest	0.47	0.39	0.41	0.48	0.44	0.44
	(0.50)	(0.49)	(0.49)	(0.50)	(0.50)	(0.50)
Avg. Days Worked Harvest	25.13	19.26	25.07	18.44	20.00	17.44
	(32.78)	(18.88)	(40.53)	(22.88)	(21.40)	(19.82)
Number of Obs.	655	598	10,088	1,001	508	10,302

Notes: Authors' calculations from LSMS-ISA data.

Table S5: Summary statistics for migrants and stayers: Uganda Waves 1-2

	(1)	(2)	(3)
	Moved-Out	Moved-In	Stayed
	mean/sd	mean/sd	mean/sd
Age	25.67	27.45	22.58
	(14.33)	(14.46)	(18.17)
% Male	0.48	0.44	0.50
	(0.50)	(0.50)	(0.50)
% Child	0.36	0.22	0.47
	(0.48)	(0.42)	(0.50)
% Other Family Member	0.41	0.54	0.44
	(0.49)	(0.50)	(0.50)
% Not a Family Member	0.03	0.08	0.00
	(0.18)	(0.27)	(0.05)
% Worked on Farm		0.29	0.40
		(0.45)	(0.49)
Number of Obs.	2,392	1,666	10,630

Notes: Authors' calculations from LSMS-ISA data.

Table S6: Summary statistics for migrants and stayers: Uganda Waves 2-3

	(1) Moved-Out mean/sd	(2) Moved-In mean/sd	(3) Stayed mean/sd
Age	25.29	26.35	21.32
	(12.61)	(13.29)	(18.43)
% Male	0.50	0.38	0.50
	(0.50)	(0.48)	(0.50)
% Child	0.33	0.23	0.50
	(0.47)	(0.42)	(0.50)
% Other Family Member	0.33	0.57	0.48
	(0.47)	(0.50)	(0.50)
% Not a Family Member	0.07	0.12	0.01
	(0.25)	(0.33)	(0.09)
% Worked on Farm	0.26	0.13	0.35
	(0.44)	(0.34)	(0.48)
Number of Obs.	1,322	482	13,733

Notes: Authors' calculations from LSMS-ISA data.

Table S7: Summary statistics for migrants and stayers: Uganda Waves 3-4

	(1) Moved-Out mean/sd	(2) Moved-In mean/sd	(3) Stayed mean/sd
Age	25.56	26.95	21.69
	(14.20)	(14.71)	(18.81)
% Male	0.46	0.45	0.49
	(0.50)	(0.50)	(0.50)
% Child	0.38	0.32	0.51
	(0.48)	(0.47)	(0.50)
% Other Family Member	0.44	0.52	0.48
	(0.50)	(0.50)	(0.50)
% Not a Family Member	0.07	0.09	0.01
	(0.25)	(0.28)	(0.08)
% Worked on Farm	0.30	0.28	0.38
	(0.46)	(0.45)	(0.48)
Number of Obs.	1,014	651	12,693

Notes: Authors' calculations from LSMS-ISA data.

Also, none of the waves of the Uganda data captures how much each person works on each plot, so each of the tables here are unable to show the average days worked for each type of migrant. Despite these data limitations, the tables show that migrants in Uganda fit a similar pattern as the other countries, and do not appear to be much different than other household members.

D Robustness and extensions

In this subsection we provide additional details for some of the robustness checks and extended analysis cited in the main text of the paper.

D.1 Instrumenting for cultivated acreage

Table S8 shows estimates similar to those in Table 3, using changed in owned acreage as an instrument for changes in cultivated acreage. Owned acreage is a strong predictor of cultivated acreage, the F-statistics for the first stage regressions are all above 118. Comparison of coefficients shows that the IV estimates are essentially the same as those in Table 3. Variation in the capacity to adjust the amount of land under cultivation is not the main driver of symmetric or asymmetric non-separation, in any of the study countries.

D.2 Excluding children from labor endowment

The main specification allows children to linearly phase-in to the household labor endowment between the ages of 11-15. To the check the robustness of that decision, we re-run the main analysis, excluding all children under the age of 15 from the household labor endowment. Therefore a household with a child reaching the age of 15 in between survey waves would increase the household's labor endowment by 1 full unit in this robustness check, whereas in the main specification, that same demographic change would increase the labor endowment by just 0.2.

The results for specifications (9)-(12) when excluding children from the labor endowment are presented in Table S9. The results are broadly consistent with our main results. The results for Tanzania provide even less evidence of asymmetry in the responses to changes in labor endowment. In Uganda, the coefficient on increases in the labor endowment is smaller than the main results, 0.182 compared to 0.267, and now only weakly significant, with a p-value of 0.054.

Table S8: Testing the separation hypothesis, instrumenting for acreage cultivated

Dependent variable:	Log of farm labor (person-days), for columns 1, 3, 5, 7 in top panel Δ Log of farm labor (person-days), all other models							
	Ethiopia		Malawi		Tanzania		Uganda	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All activities								
Labor endow. (E)	0.564***		0.469***		0.679***		0.304***	
	(0.000)		(0.000)		(0.000)		(0.000)	
ΔE		0.616***		0.542***		0.699***		0.214***
		(0.000)		(0.000)		(0.000)		(0.001)
Observations	5,652	2,825	4,818	2,095	5,901	3,895	8,218	5,362
R^2	0.558	0.088	0.425	0.101	0.251	0.025	0.275	0.089
By activity								
Reference $\times \Delta E$	0.476***		0.589***		0.646***		0.214***	
	(0.004)		(0.000)		(0.000)		(0.001)	
Reference $\times \Delta E^+$		0.185		0.834***		0.667***		0.274***
		(0.443)		(0.000)		(0.002)		(0.007)
Reference $\times \Delta E^-$		0.701***		0.202		0.628**		0.159*
		(0.006)		(0.421)		(0.011)		(0.084)
Harvest $\times \Delta E$	0.449***		0.287**		0.591***			
	(0.007)		(0.029)		(0.000)			
Harvest $\times \Delta E^+$		0.183		0.355*		0.760***		
		(0.473)		(0.062)		(0.000)		
Harvest $\times \Delta E^-$		0.654**		0.182		0.430*		
		(0.019)		(0.506)		(0.074)		
Weeding $\times \Delta E$					0.768***			
					(0.000)			
Weeding $\times \Delta E^+$						0.925***		
						(0.000)		
Weeding $\times \Delta E^-$						0.619**		
						(0.011)		
Reference: Inc = Dec	.	0.169	.	0.071	.	0.915	.	0.437
Harvest: Inc = Dec	.	0.265	.	0.649	.	0.351	.	.
Weeding: Inc = Dec	.	.	.	.	.	0.388	.	.
Reference = Harvest	0.792	.	0.001	.	0.477	.	.	.
Reference = Weeding	.	.	.	.	0.102	.	.	.
Weeding = Harvest	.	.	.	.	0.011	.	.	.
Reference = Harvest, +	.	0.992	.	0.001	.	0.487	.	.
Reference = Harvest, −	.	0.764	.	0.871	.	0.105	.	.
Reference = Weeding, +	.	.	.	.	.	0.035	.	.
Reference = Weeding, −	.	.	.	.	.	0.940	.	.
Weeding = Harvest, +	.	.	.	.	.	0.115	.	.
Weeding = Harvest, −	.	.	.	.	.	0.056	.	.
Observations	5,650	5,650	4,190	4,190	11,685	11,685	5,362	5,362
R^2	0.028	0.028	0.149	0.150	0.028	0.028	0.089	0.089

Table S9: Testing the separation hypothesis, excluding children < 15 from LE

Dependent variable:	Log of farm labor (person-days), for columns 1, 3, 5, 7 in top panel Δ Log of farm labor (person-days), all other models							
	Ethiopia		Malawi		Tanzania		Uganda	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All activities								
Labor endow. (E)	0.630***		0.562***		0.640***		0.302***	
	(0.000)		(0.000)		(0.000)		(0.000)	
ΔE		0.660***		0.581***		0.605***		0.159**
		(0.000)		(0.000)		(0.000)		(0.015)
Observations	5,652	2,825	4,818	2,095	5,901	3,895	8,218	5,362
R^2	0.559	0.089	0.433	0.104	0.249	0.023	0.273	0.089
By activity								
Reference $\times \Delta E$	0.563***		0.589***		0.603***		0.159**	
	(0.000)		(0.000)		(0.000)		(0.015)	
Reference $\times \Delta E^+$		0.307		0.784***		0.559***		0.182*
		(0.150)		(0.000)		(0.006)		(0.054)
Reference $\times \Delta E^-$		0.763***		0.267		0.653***		0.138
		(0.001)		(0.264)		(0.004)		(0.123)
Harvest $\times \Delta E$	0.513***		0.340***		0.550***			
	(0.002)		(0.009)		(0.000)			
Harvest $\times \Delta E^+$		0.195		0.374**		0.623***		
		(0.372)		(0.026)		(0.002)		
Harvest $\times \Delta E^-$		0.770***		0.299		0.471**		
		(0.003)		(0.237)		(0.026)		
Weeding $\times \Delta E$					0.687***			
					(0.000)			
Weeding $\times \Delta E^+$						0.716***		
						(0.000)		
Weeding $\times \Delta E^-$						0.656***		
						(0.002)		
Reference: Inc = Dec	.	0.155	.	0.107	.	0.777	.	0.732
Harvest: Inc = Dec	.	0.105	.	0.817	.	0.614	.	.
Weeding: Inc = Dec	.	.	.	.	.	0.847	.	.
Reference = Harvest	0.624	.	0.003	.	0.443	.	.	.
Reference = Weeding	.	.	.	.	0.215	.	.	.
Weeding = Harvest	.	.	.	.	0.024	.	.	.
Reference = Harvest, +	.	0.491	.	0.001	.	0.569	.	.
Reference = Harvest, −	.	0.965	.	0.790	.	0.130	.	.
Reference = Weeding, +	.	.	.	.	.	0.116	.	.
Reference = Weeding, −	.	.	.	.	.	0.978	.	.
Weeding = Harvest, +	.	.	.	.	.	0.328	.	.
Weeding = Harvest, −	.	.	.	.	.	0.034	.	.
Observations	5,650	5,650	4,190	4,190	11,685	11,685	5,362	5,362
R^2	0.029	0.030	0.152	0.153	0.027	0.027	0.089	0.089

Notes: Authors' calculations from LSMS-ISA data. Top panel, odd-numbered columns, are pooled regressions, with enumeration area fixed effects. Bottom panel and top panel, even-numbered columns, are household fixed effects regressions implemented in first differences. Estimates are marginal effects, interpretable as elasticities. p-values in parentheses. Significance: *** 0.01, ** 0.05, * 0.1. Standard errors clustered at the enumeration area. In all regressions, changes in acreage owned by the household is used as an instrument for changes in cultivated acreage. All regressions include controls for changes in demographic shares, and year fixed effects. The "Reference" activity is "All non-harvest" for Ethiopia and Malawi, "Planting and land preparation" for Tanzania, and "All farming activities" for Uganda. Labor endowment LE is measured in logs in all specifications. ΔLE^+ is the change in labor endowment if positive, and 0 otherwise; ΔLE^- is the change in labor endowment if negative, and 0 otherwise. The dependent variable is defined so that everyone over age 15 is 1 worker, and children aged 11-14 count as $(Age - 10) \times 0.2$ workers.

D.3 Heterogeneous effects by agro-ecological zone

As discussed above in Section 2.5, the labor market may operate differently in different agro-ecological zones (AEZs). The LSMS-ISA data identify the relevant AEZ for each household. For each country, two AEZs cover the majority of households. To preserve space, we then group all other AEZs together as a third grouping. Then for each country, we interact all **F** and ΔE terms with dummy variables for each AEZ.

Table S10: Testing the separation hypothesis across agro-ecological zones in Ethiopia

	Tropic-cool/semiarid		Tropic-cool/subhumid		Other AEZs	
	Sym	Asym	Sym	Asym	Sym	Asym
	(1)	(2)	(3)	(4)	(5)	(6)
Reference $\times\Delta E$	0.947***		0.224		0.357	
	(0.007)		(0.350)		(0.155)	
Reference $\times\Delta E^{+}$		0.947*		0.114		-0.330
		(0.063)		(0.767)		(0.279)
Reference $\times\Delta E^{-}$		0.952*		0.299		0.908**
		(0.081)		(0.373)		(0.042)
Harvest $\times\Delta E$	0.786**		0.356		0.249	
	(0.017)		(0.107)		(0.404)	
Harvest $\times\Delta E^{+}$		0.507		0.047		0.062
		(0.350)		(0.902)		(0.878)
Harvest $\times\Delta E^{-}$		0.998*		0.580		0.424
		(0.056)		(0.109)		(0.426)
Reference: Inc = Dec	.	0.995	.	0.731	.	0.037
Harvest: Inc = Dec	.	0.558	.	0.369	.	0.625
Reference = Harvest	0.393	.	0.331	.	0.590	.
Reference = Harvest, +	.	0.094	.	0.803	.	0.275
Reference = Harvest, −	.	0.878	.	0.207	.	0.087
Observations	5,650	5,650	.	.	.	.
R^2	0.032	0.033	.	.	.	.

The results for Ethiopia are presented in Table S10. The "Tropic-cool / semiarid" zone covers 31.4% of the households and the "Tropic-cool / subhumid" zone covers another 40.0% of households. The remaining 28.6% of households are spread across six other AEZs, and are grouped together for this analysis. For clarity, the results for each AEZ are grouped into their own column. So, in Table S10, the results shown in the odd-numbered columns are from one regression, based on specification (11), and all the even-numbered columns are from one regression based on specification (12). Then columns (1) and (2) show the results for the "Tropic-cool / semiarid" zone, columns (3) and (4) show the results for the "Tropic-cool/subhumid" zone, and then the last two columns are for the other AEZs. The results for the "Tropic-cool / semiarid" zone show the strongest evidence of non-separation. However the pattern supporting a labor supply constraint is only evident during the harvest season. The results for the "Tropic-cool / subhumid" zone do not show any evidence of non-separation.

Table S11: Testing the separation hypothesis across agro-ecological zones in Malawi

	Tropic-warm/semiarid		Tropic-warm/subhumid		Other AEZs	
	Sym	Asym	Sym	Asym	Sym	Asym
	(1)	(2)	(3)	(4)	(5)	(6)
Reference $\times \Delta E$	0.284		0.864***		0.849***	
	(0.100)		(0.000)		(0.001)	
Reference $\times \Delta E^{+}$		0.433*		1.092***		1.009***
		(0.094)		(0.000)		(0.004)
Reference $\times \Delta E^{-}$		-0.022		0.623		0.662
		(0.950)		(0.163)		(0.164)
Harvest $\times \Delta E$	0.135		0.283		0.870***	
	(0.485)		(0.185)		(0.001)	
Harvest $\times \Delta E^{+}$		0.256		0.098		1.030***
		(0.378)		(0.755)		(0.001)
Harvest $\times \Delta E^{-}$		0.064		0.451		0.473
		(0.872)		(0.310)		(0.344)
Reference: Inc = Dec	.	0.358	.	0.425	.	0.587
Harvest: Inc = Dec	.	0.734	.	0.567	.	0.372
Reference = Harvest	0.300	.	0.000	.	0.904	.
Reference = Harvest, +	.	0.460	.	0.000	.	0.923
Reference = Harvest, −	.	0.592	.	0.399	.	0.618
Observations	4,190	4,190	.	.	.	.
R^2	0.168	0.168	.	.	.	.

The results for Malawi are shown in Table S11. The "Tropic-warm / semiarid" zone covers 46.7% of the households and the "Tropic-warm / subhumid" zone covers another 30.9% of households. The remaining 22.4% of households are spread across two "Tropic-cool" zones. The results in columns (1) and (2) of Table S11, for the "Tropic-warm / semiarid" zone show no evidence of non-separation. Whereas the "Tropic-warm / subhumid" zone only shows evidence of non-separation during cultivation, with results showing the pattern of a labor surplus as in the main analysis. The results for the two "Tropic-cool" zones show evidence of non-separation in both seasons, and have results consistent with a binding labor demand constraint.

The results for Tanzania are shown in Table S12. The "Tropic-warm / subhumid" zone covers 55.9% of households and the "Tropic-cool / subhumid" zone covers another

Table S12: Testing the separation hypothesis across agro-ecological zones in Tanzania

	Tropic-warm/subhumid		Tropic-cool/subhumid		Other AEZs	
	Sym	Asym	Sym	Asym	Sym	Asym
	(1)	(2)	(3)	(4)	(5)	(6)
Reference $\times \Delta E$	0.788***		0.189		0.862***	
	(0.000)		(0.445)		(0.006)	
Reference $\times \Delta E^{+}$		1.127***		-0.462		0.472
		(0.000)		(0.231)		(0.289)
Reference $\times \Delta E^{-}$		0.487		0.777*		1.208*
		(0.137)		(0.070)		(0.063)
Harvest $\times \Delta E$	0.778***		0.139		0.576*	
	(0.000)		(0.587)		(0.089)	
Harvest $\times \Delta E^{+}$		1.119***		0.208		0.115
		(0.000)		(0.586)		(0.750)
Harvest $\times \Delta E^{-}$		0.438		0.115		1.003*
		(0.171)		(0.756)		(0.099)
Weeding $\times \Delta E$	1.020***		0.284		0.507*	
	(0.000)		(0.245)		(0.092)	
Weeding $\times \Delta E^{+}$		1.283***		0.175		0.601
		(0.000)		(0.666)		(0.151)
Weeding $\times \Delta E^{-}$		0.752**		0.416		0.448
		(0.024)		(0.281)		(0.392)
Reference: Inc = Dec	.	0.185	.	0.059	.	0.429
Harvest: Inc = Dec	.	0.149	.	0.867	.	0.232
Weeding: Inc = Dec	.	0.263	.	0.699	.	0.835
Reference = Harvest	0.918	.	0.750	.	0.158	.
Reference = Weeding	0.007	.	0.367	.	0.207	.
Weeding = Harvest	0.003	.	0.258	.	0.778	.
Reference = Harvest, +	.	0.959	.	0.003	.	0.366
Reference = Harvest, −	.	0.745	.	0.012	.	0.517
Reference = Weeding, +	.	0.306	.	0.003	.	0.538
Reference = Weeding, −	.	0.061	.	0.143	.	0.114
Weeding = Harvest, +	.	0.195	.	0.866	.	0.212
Weeding = Harvest, −	.	0.008	.	0.091	.	0.077
Observations	11,685	11,685	.	.	.	.
R^2	0.030	0.031	.	.	.	.

28.7%. The remaining 15.5% of households are located in four other AEZs. The results in columns (1) and (2) for the "Tropic-warm / subhumid" zone show strong evidence of non-separation, with results consistent with a binding labor demand constraint during the planting and weeding seasons. The results for the "Tropic-cool / subhumid" region do not show any evidence of non-separation, while the other AEZs exhibit results consistent with a binding labor supply constraint during planting and weeding.

The results for Uganda are presented in Table S13. The "Tropic-warm / humid" zone covers 49.8% of households and the "Tropic-cool / humid" zone covers another 29.5%

of households. The remaining 20.8% of households are located in two "subhumid" zones. The results in columns (1) and (2) for the "Tropic-warm / humid" zone are consistent with the main results for Uganda, that of a binding labor demand constraint. The results for the "Tropic-cool / humid" zone do not reveal any evidence of non-separation. The results in the last two columns for the "subhumid" zones also show some weak support for non-separation due to a binding labor demand constraint.

Table S13: Testing the separation hypothesis across agro-ecological zones in Uganda

	Tropic-warm/humid		Tropic-cool/humid		Other AEZs	
	Sym	Asym	Sym	Asym	Sym	Asym
	(1)	(2)	(3)	(4)	(5)	(6)
Reference $\times\Delta E$	0.247***		0.103		0.236*	
	(0.004)		(0.380)		(0.093)	
Reference $\times\Delta E^{+}$		0.303**		0.080		0.425*
		(0.023)		(0.646)		(0.072)
Reference $\times\Delta E^{-}$		0.199		0.126		0.001
		(0.118)		(0.452)		(0.997)
Reference: Inc = Dec	.	0.594	.	0.853	.	0.222
Observations	5,362	5,362	.	.	.	.
R^2	0.091	0.091	.	.	.	.

D.4 Heterogeneous effects by wealth

As discussed in Section 2.3.1, it is possible for a failure in the credit market to manifest appear as a binding labor supply constraint in our analysis. We test for this explanation in Ethiopia, as Ethiopia is the only country which exhibited a binding labor supply constraint in the main analysis. The results by wealth groups for Ethiopia are shown in Table 5.

However, it may also be the case that wealthy households are able to respond to changes in their household labor endowment in more efficient ways than poor households. Wealthy households may have more access to credit to help hire more workers to replace a household member which moved out, or they may have better access to land to make use of new household members. To test for heterogeneity in the separation hypothesis along the wealth dimension, we use two measures of household wealth, expenditure-per-capita and an asset index. We then create a dummy variable for whether each household had above median wealth in the first period of the LSMS-ISA data (waves 1 for Malawi and Tanzania, but wave 2 for Uganda), and interact that dummy variable with all $\mathbf{F}$ and ΔE terms in specifications (11) and (12). The results of these analyses are shown in Table S14 for Malawi, Tanzania, and Uganda.

The results for Malawi are shown in columns (1) - (4). For both measures of wealth, the results show evidence of a binding labor demand constraint for wealthy households in Malawi. The excess labor for wealthy households is evident during both cultivation and harvest. The poor households show evidence of non-separation during cultivation, and during harvest only when using the asset measure. The asymmetric results for Malawi indicate a binding labor demand constraint for poor households only when using the expenditure-per-capita measure for wealth.

The Tanzania results are shown in columns (5) - (8). The results for the symmetric analysis finds evidence of non-separation in all but one agricultural phases, for both wealthy and poor households, using both measures of wealth. When using the asset index measure of wealth, the wealth households exhibit a pattern of a binding labor demand constraint, however the results when using the expenditure-per-capita measure of wealth show significant results on both increases and decreases in the household labor endowment. Households in

Tanzania which have an expenditure-per-capita below the median show some evidence of a binding labor demand constraint, but that pattern is not upheld when using the asset index.

For Uganda, the results are shown in columns (9) - (12). The results mostly show evidence of a binding labor demand constraint for both wealthy and poor households using both measures of wealth. However, the coefficient for poor households according to the asset index is not statistically significant.

Table S14: Testing the separation hypothesis across wealth groups in Malawi, Tanzania, and Uganda

	Malawi				Tanzania				Uganda			
	Exp PC		Assets		Exp PC		Assets		Exp PC		Assets	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wealth $< 50\%$												
Reference $\times \Delta E$	0.502***		0.737***		0.556***		0.421**		0.186**		0.086	
	(0.003)		(0.000)		(0.008)		(0.042)		(0.030)		(0.751)	
Reference $\times \Delta E^{+}$		0.552**		0.922***		0.789**		0.288		0.269**		0.232
		(0.035)		(0.000)		(0.022)		(0.380)		(0.034)		(0.685)
Reference $\times \Delta E^{-}$		0.435		0.364		0.294		0.547*		0.102		-0.051
		(0.242)		(0.280)		(0.460)		(0.064)		(0.463)		(0.866)
Harvest $\times \Delta E$	0.151		0.408**		0.338		0.429**					
	(0.405)		(0.028)		(0.118)		(0.034)					
Harvest $\times \Delta E^{+}$		-0.032		0.389		0.591*		0.440				
		(0.914)		(0.135)		(0.082)		(0.139)				
Harvest $\times \Delta E^{-}$		0.408		0.446		0.055		0.419				
		(0.250)		(0.232)		(0.883)		(0.154)				
Weeding $\times \Delta E$					0.563***		0.640***					
					(0.006)		(0.001)					
Weeding $\times \Delta E^{+}$						0.881***		0.640**				
						(0.007)		(0.031)				
Weeding $\times \Delta E^{-}$						0.207		0.640**				
						(0.578)		(0.032)				
Reference: Inc = Dec	.	0.826	.	0.219	.	0.419	.	0.579	.	0.407	.	0.701
Harvest: Inc = Dec	.	0.418	.	0.911	.	0.342	.	0.960	.	.	.	.
Weeding: Inc = Dec	.	.	.	.	.	0.233	.	0.999	.	.	.	.
Wealth $\geq 50\%$												
Reference $\times \Delta E$	0.733***		0.498**		0.727***		0.955***		0.221**		0.216***	
	(0.000)		(0.015)		(0.001)		(0.000)		(0.013)		(0.001)	
Reference $\times \Delta E^{+}$		1.120***		0.794**		0.541*		1.173***		0.265*		0.269***
		(0.000)		(0.012)		(0.096)		(0.000)		(0.086)		(0.005)
Reference $\times \Delta E^{-}$		0.034		0.149		0.888***		0.739*		0.185		0.168
		(0.913)		(0.665)		(0.005)		(0.059)		(0.124)		(0.104)
Harvest $\times \Delta E$	0.503***		0.308		0.852***		0.804***					
	(0.005)		(0.127)		(0.000)		(0.000)					
Harvest $\times \Delta E^{+}$		0.802***		0.512*		0.996***		1.151***				
		(0.001)		(0.075)		(0.001)		(0.000)				
Harvest $\times \Delta E^{-}$		-0.035		0.068		0.735**		0.461				
		(0.923)		(0.857)		(0.014)		(0.231)				
Weeding $\times \Delta E$					0.981***		0.961***					
					(0.000)		(0.000)					
Weeding $\times \Delta E^{+}$						1.019***		1.312***				
						(0.000)		(0.000)				
Weeding $\times \Delta E^{-}$						0.951***		0.612				
						(0.002)		(0.100)				
Reference: Inc = Dec	.	0.019	.	0.216	.	0.470	.	0.432	.	0.702	.	0.493
Harvest: Inc = Dec	.	0.074	.	0.401	.	0.560	.	0.183	.	.	.	.
Weeding: Inc = Dec	.	.	.	.	.	0.877	.	0.174	.	.	.	.
Ref.: Top = Bottom	0.299	.	0.376	.	0.576	.	0.074	.	0.768	.	0.646	.
Ref.: Top = Bottom, +	.	0.117	.	0.742	.	0.627	.	0.052	.	0.981	.	0.950
Ref.: Top = Bottom, −	.	0.397	.	0.648	.	0.245	.	0.685	.	0.641	.	0.512
Harvest: Top = Bottom	0.160	.	0.723	.	0.080	.	0.211	.	.	.	.	.
Harvest: Top = Bottom, +	.	0.033	.	0.754	.	0.398	.	0.069	.	.	.	.
Harvest: Top = Bottom, −	.	0.342	.	0.471	.	0.136	.	0.929	.	.	.	.
Weeding: Top = Bottom	.	.	.	.	0.080	.	0.211	.	.	.	.	.
Weeding: Top = Bottom, +	.	.	.	.	.	0.398	.	0.069	.	.	.	.
Weeding: Top = Bottom, −	.	.	.	.	.	0.136	.	0.929	.	.	.	.
Observations	4,190	4,190	4,190	4,190	11,685	11,685	11,685	11,685	5,362	5,362	5,362	5,362
R^2	0.157	0.159	0.158	0.159	0.030	0.030	0.032	0.032	0.090	0.090	0.090	0.090