Land Consolidation and Rural Labor Markets: Theory and Evidence From Colombia

Karen Ortiz-Becerra[†]
UC Davis

November 1, 2021

Latest version available here

Abstract

The consolidation of land in large farms is accelerating in many developing countries. This paper studies the implications of this consolidation on rural employment and workers' welfare. I develop a general equilibrium model of local labor markets that considers both farm and nonfarm labor. Large-farm consolidation affects the demand for farm labor via labor intensity and the nonfarm labor demand through nonhomothetic consumption growth. The model shows that consolidation may reduce workers' income if the pull response in the nonfarm sector is small relative to the shift in farm labor demand. Using quasi-experimental variation and a novel dataset from Colombia, I find that large-farm consolidation decreased the share of agricultural labor and led to a fourteen percentage point increase in the unemployment rate. These findings shed light on the distributional impacts of consolidation across individuals and the spatial implications of the reallocation of labor out of the agricultural sector.

[†]PhD Candidate in Agricultural and Resource Economics, University of California Davis. E-mail: kaortizb@ucdavis.edu. I am deeply grateful to Ashish Shenoy, Michael Carter, Dalia Ghanem, and Steve Boucher for providing guidance and advising on this project and throughout my degree. Many thanks to Douglas Gollin, Diego Restuccia, Jason Shogren, Gaurav Chiplunkar, Esteban Quiñonez, Ameet Morjaria, Debi Mohapatra, John Morehouse, Eleanor Wiseman, Fabio Sanchez, Margarita Gafaro, Jessica Rudder, Aleksander Michuda, and seminar audiences at MWIEDC, UC Davis, the 1st Workshop in Applied Microeconomics, and the LACEA-RIDGE Job Market Showcase for helpful discussions and providing valuable feedback. I also thank Olam International, its team of in-country operators, and the farmers in rural regions of Antioquia, Huila, Casanare, and Meta for their time during my fieldwork visits throughout the summer and fall of 2019. I acknowledge the support from the UC Davis Provost's Dissertation Fellowship, the Bert and Nell Krantz Fellowship, and the Henry A. Jastro Graduate Research Award.

1 Introduction

The consolidation of land in extensive farms is accelerating in developing countries. In the past two decades, more than four thousand large-scale deals have taken place. Close to 60% of these deals are transactions of at least five thousand hectares, and the median size across all transactions is seven thousand hectares.[1] Scholars suggest that this rise in demand for agricultural land is unlikely to slow given population growth trends (Liao et al., 2020; Deininger et al., 2011). Therefore, it is crucial to understand the implications of this consolidation on economic development, especially in the rural economies where this consolidation occurs.

Large-scale production has the potential to increase aggregate agricultural productivity. Although technology usually exhibits constant returns to scale, market imperfections often render production at scale more efficient (Ma et al., 2021; Foster and Rosenzweig, 2017; Collier and Dercon, 2014; Chavas, 2001; Putterman, 1983). Large producers have better access to financial and output markets and the ability to benefit from pecuniary economies in the procurement of inputs. These favorable conditions usually outweigh the advantages of small producers in access to local knowledge and labor supervision costs.[2] The productivity gains of land consolidation, however, are not necessarily evenly distributed across individuals. Labor is an essential endowment for a large proportion of rural households, and the extent to which these populations benefit from such consolidation depends on local employment gains.

In this paper, I study the implications of large-farm consolidation on employment in rural economies. I focus on three main questions of interest: Does the consolidation of land affect the local allocation of labor across sectors? Do rural workers benefit from this consolidation? And what are the potential economic mechanisms driving the effects?[3]

I study these questions in the context of rural Colombia, where land concentration has been historically high, and 40% to 50% of agricultural land is consolidated in large farms. Two main features make this setting particularly relevant to study these questions. First, a large majority of rural households are landless (58%) and derive most of their income from labor markets. Second, one of the provisions of the 2016 Peace Agreement is to redis-

[1] These numbers refer to completed transactions of at least five hundred hectares for agricultural activities. For more details on this data, see the Land Matrix Project.

[2] Of course, some crops can still be produced competitively at different scales depending on local factor endowments and labor costs (Deininger et al., 2011). There are also certain crops that are often more profitable at small scale such as specialty and certified coffee.

[3] Large-scale transactions can also have political economy implications for rural economies since they are more prevalent in countries with weak security of land rights (Deininger, 2011; Davis et al., 2014). While these concerns merit a thorough understanding of these implications, such analysis is outside of the scope of this paper.

tribute land to landless peasants and farmers with insufficient acreage. Recent estimates suggest that at least 2.6 million hectares are available to implement this redistribution policy (Arteaga et al., 2017). Many of the stylized patterns that I establish in the Colombian setting are likely general to rural economies in other developing contexts. Thus, this analysis is also relevant for other low and middle-income countries where large-scale land deals are taking place.

This paper develops a two-sector general equilibrium model of local economies to analyze the implications of land consolidation on employment and the equilibrium wage. This model combines two microeconomic insights that are ordinary in previous agricultural production models – differences in factor intensity across the farm size distribution and efficiency gains from scale – with a third feature usual in rural economies that has not yet been formalized – the non-homothetic demand for local goods. A key feature of the model is that land consolidation affects the demand for labor in both economic sectors. On the one hand, labor intensity impacts the demand for farm labor. On the other, nonfarm labor demand is affected by non-homothetic consumption growth.

Using this framework, I show that consolidation leads to a push response in the demand for labor in the farm sector and a likely pull response in nonfarm labor demand. In consequence, if the pull response is small relative to the push response in the farm sector, consolidation may lead to a reallocation of workers out of the agricultural sector along with a reduction in wages. These theoretical results highlight the importance of local multiplier effects to increase labor demand and attain gains in worker's income, which is empirically relevant in developing contexts where labor mobility barriers reduce workers' reallocation across space (Morten and Oliveira, 2018; Dix-Carneiro and Kovak, 2017; Bryan et al., 2014).

Next, I examine the impact of a specific shift in land concentration that occurred during the 1990s. This shift was largely driven by a change in the terms of trade that transformed land use in the country. The area in pastures and land-intensive crops increased by more than two million hectares, while coffee and other labor-intensive crops lost important participation in agricultural land (Balcázar, 2003). I study this shift in land concentration using a novel panel dataset of rural municipalities with measures on employment and large-farm consolidation. I examine impacts on the reallocation of labor across sectors and workers' income, proxied by measures of wages and unemployment rates.

To estimate treatment effects, I leverage quasi-experimental variation in the ability of local economies to respond to the change in the terms of trade. Specifically, I construct an instrument for large-farm consolidation based on topographic features that influence the suitability to produce a given crop at a large scale. The intuition behind this instrument is that physical and financial constraints render flatter terrains more suitable for production at scale. I account for two main confounders that might be correlated with a town's topography. First, labor market trends induced by the change in terms of trade that are associated with crop suitability (e.g. crop-specific income effects due to changes in relative prices). Second, the upsurge in conflict in rural areas, which had potential differential effects across economies in the lowlands and the slopes. In this analysis, identification comes from the variation in the physical ability to produce a certain crop a scale, rather than the suitability to produce a particular crop.

The findings indicate that the consolidation of land in large farms prompted the reallocation of labor out of agriculture within local economies. Specifically, one standard deviation increase in consolidation resulted in a 10 percentage point decrease in the share of agricultural workers.The results further suggest that this reallocation of labor out of the agricultural sector did not bring improvements in workers' income. While there is no detectable significant effect on rural wages, I find that consolidation led to a 14 percentage point increase in the unemployment rate. Taken together, these findings indicate that consolidation led to a decrease in the demand for farm workers and an absolute reduction in the demand for rural labor. These results are consistent with the model's predictions when the pull response in the nonfarm sector is small relative to the push response in the farms.

I provide evidence that these findings are not driven by several factors that may confound the relationship between employment and the instrument. First, parallel pre-trends support the exclusion restriction that topography does not lead to differential trends in labor markets across similar towns. Second, results are unchanged when controlling for the upsurge of conflict and forced displacement, a major source of rural migration during this period. These findings are also robust to the use of different inference procedures, including alternative types of intra-cluster correlation and size corrections for weak instruments.

Overall, these findings provide evidence that large-farm consolidation may lower the welfare of rural workers. These results contrast with previous work that focuses on the gains in aggregate productivity and shed light on the distributional impacts of consolidation across individuals. My findings also indicate that the reallocation of labor out of agriculture in rural economies may be accompanied by decrease in workers' income, which is consistent with recent evidence on the spatial impacts of structural transformation in the presence of barriers to labor mobility (Eckert and Peters, 2018).

The theoretical framework developed in this paper has implications for the design of future land policies. The model suggests that the effect on the equilibrium wage depends on the initial level of consolidation since the strength of the push and pull responses in labor demand in each sector varies across the farm size distribution. Similarly, the effect on the wage also depends on the type of consolidation that takes place. In particular, whether land is consolidated into exceedingly large farms or medium-size farms where local multiplier effects are likely higher. In work in progress, I calibrate the model to Colombian farm-level data to examine two counterfactual scenarios that might take place with the future redistribution policy.[4] One redistributes land from large to middle-sized farms while the other increases the proportion of area in small farms. For each counterfactual, I will quantify the effects on the labor reallocation across sectors and the equilibrium wage under different extents of labor mobility across space. In addition, I will assess how much of these effects are driven by non-homothetic local consumption growth.

Related literature: This paper contributes to the literature that examines the relationship between land allocation and economic development. This literature includes theoretical work that micro-founds the connections between scale and aggregate productivity (Ma et al., 2021; Henderson and Isaac, 2017; Eswaran and Kotwal, 1986; Carter and Kalfayan, 1989), macroeconomic models that quantify the contribution of misallocation to gaps in productivity (Santaeulàlia-Llopis, 2021; Adamopoulos and Restuccia, 2020; Adamopoulos et al., 2019; Chen, 2017; Adamopoulos and Restuccia, 2014), and more recent empirical work that examines the effects of large-scale transactions on rural populations beyond productivity gains (Liao et al., 2020; Ali et al., 2019; Deininger and Xia, 2016). My work focuses on implications for rural labor markets and adds to this literature in several ways. First, it introduces a framework that considers the connection between income concentration on labor demand in the nonfarm sector. Most existing models only consider labor implications in the farm sector despite the previous discussion of this connection (Mellor, 2017; Ranis and Stewart, 1993). This framework also incorporates insights from the literature on the rural nonfarm economy, such as the importance of local consumption on employment in this sector (Foster, 2011; Haggblade et al., 2009, 2007; Foster and Rosenzweig, 2008; Lanjouw and Lanjouw, 2001). By adding these new insights, I show that land consolidation might negatively impact workers' earnings, even when there are scale economies in production.

In addition to the theoretical framework, this paper provides novel empirical evidence on the net effects of large-scale consolidation on rural labor markets. A growing number of papers examine the effects of large-scale land investments on a wide set of outcomes. However, due to data constraints, many conduct descriptive work or rely on cross-sectional data and the unconfoundedness assumption to identify treatment effects (Bottazzi et al., 2018; Herrmann, 2017; Jiao et al., 2015). Two important exceptions are Ali et al. (2019) and

[4]See Appendix B for more details on this ongoing analysis.

Deininger and Xia (2016), who exploit data on the change in the exposure of communities to large-scale investments. Yet, since their aim is to estimate spillover effects on smallholders' investments, these analyses do not provide evidence on the economy-wide impacts on employment and wages. By focusing on the aggregate net effect on the set of landed and landless individuals, this paper provides a comprehensive picture of the implications of land consolidation for labor markets.

A second line of work related to this paper is the literature on structural transformation (Herrendorf et al., 2014). My findings show that land consolidation is another driver of labor reallocation across sectors, and that a shift in labor out of agriculture can occur along with a decrease in wages. This work is also related to recent studies that examine the spatial implications of structural transformation (Eckert and Peters, 2018; Bustos et al., 2016; Nagy, 2016; Desmet and Rossi-Hansberg, 2014; Michaels et al., 2012; Caselli and Coleman II, 2001). My results are consistent with the analysis by Eckert and Peters (2018) who quantify a reduction in rural income per capita after the transformation of the US economy during the 20th century.

The rest of this paper is organized as follows. In the next section, I document empirical patterns that relate large-farm consolidation and employment in each sector. Section 3 develops the model and discuss its main analytical insights. Section 4 examines the impact of large-farm consolidation during the 1990s on rural labor markets in Colombia, and Section 5 concludes.[5]

2 Empirical Patterns Relating Farm Size and Labor Demand

In this section, I establish four descriptive patterns that relate farm size to the demand for workers in agricultural and nonagricultural sectors. These patterns motivate the mechanisms and the structure of the model that I propose in Section 3. These patterns are obtained using data from the *2014 National Agricultural Census*, the *2016 National Household Survey*, and the cost study of agricultural crops conducted by Perfetti et al. (2012).

Pattern 1. *Input intensity in agriculture varies with farm size*

One empirical regularity that has been documented in low and middle-income countries is that labor intensity per hectare decreases with farm size (Carter, 1984; Sen, 1981; Deininger et al., 2016). Figure 1 shows that this relationship holds for the main crops in Colombia, including coffee, oil palm, and rice.[6] By contrast, Figure 2 shows that large

[5]For details on the ongoing calibration exercise, see Appendix B.

[6]This relation remains practically unchanged when using the total area cultivated as opposed to the total

producers are more likely to use machinery and small equipment tools for agricultural production. In particular, farmers with 500+ hectares are 5.5 times more likely to own equipment than farmers with less than 3 hectares of land. This variation in input intensity across the size distribution implies that large producers are more likely to substitute labor for farm equipment in production. Thus, a farmer's demand for workers per hectare will depend her the farm size.

Pattern 2. *Nonfarm employment depends largely on local consumption*

By 2017, 40% of rural workers had a primary job in the nonagricultural sector. This sector is composed of all activities that do not make an intensive use of land to raise livestock or produce crops and raw materials. In Figure 3, I show that two-thirds of the workers in this sector work in industries that produce services or goods that are nontradable. Notably, 19% work in the provision of personal and social services, 8% work on construction, and 22.5% work in retail.[7] These observations imply that local consumption is a crucial driver of employment in the nonfarm sector. The following patterns suggest two ways through which this consumption could be affected by the concentration of land in large farms.

Pattern 3. *Large producers benefit from economies of scale in equipment markets*

Figure 4 documents that large producers usually spend less money per hectare on farm equipment. This regularity holds for different crops and across multiple agricultural regions in the country. Given the patterns in equipment use documented above, this difference in costs implies economies of scale in equipment markets. In particular, large producers tend to face a lower rental price of farm equipment than their smaller counterparts.[8] This observation is further supported by qualitative interviews in the field that point to the existence of nonlinear pricing in the cost of machinery services. For instance, the rental price of a tractor service per hectare is U$40 for farms with less than 100 hectares and U$33 (or less) for larger farms.[9] These advantages in equipment markets render production at scale more profitable. Thus, to the extent that these profits are used for local consumption,

farm size. It also holds when using the total number of permanent workers as a measure of employment instead of the total man days.

[7] I consider retail goods as nontradable, even though some of these are imported from larger cities. The motivation behind this choice is that these goods are almost exclusively consumed within the local economies.

[8] Formally, let $\frac{p_j k_j}{h_j}$ be the the total cost of equipment per hectare for a farm with size j, where $j =$ small, large. If $\frac{k_s}{h_s} < \frac{k_l}{h_l}$, then $\frac{p_s k_s}{h_s} > \frac{p_s k_s}{h_s}$ implies that $\frac{p_s}{p_l} > \frac{k_l h_s}{k_s h_l} > 1$.

[9] These nonlinearities in input pricing are not unique to equipment markets. Bulk pricing is a common practice in the markets for fertilizers and pesticides.

large producers can generate higher multipliers effects in the nonfarm economy.

Pattern 4. *Large producers are less likely to spend a high share of their income in the local economy*

Multiple scholars have documented the prevalence of absent landholders in Colombia (Adams, 1966; Edel, 1971; Piniero, 2016). These landholders are usually large producers that live in the city most of the year. In recent work, Robineau et al. (2010) document that 40% of the farmland in two municipalities close to the capital belongs to absentee owners who hire a manager to take care of the production. More generally, the *2014 Agrarian Census* reports that a large share of producers (73.3%) do not live permanently on their farms. One implication of the absenteeism of large producers is that the demand for nontradable goods in the rural economy will depend on how concentrated the agricultural profits are. This is also the case when all producers live in the rural economies, as it is reasonable to think that wealthier producers are more likely to replace local goods with higher-quality counterparts produced and sold in the cities (De Janvry and Sadoulet, 1993).

3 Two-Sector Model of Local Labor Markets

In this section, I develop a static model of a rural economy with two sectors. This model features a small-open economy where farmers are heterogeneous in their landholdings, input demands vary with plot size, and local consumption is crucial for rural employment. This model aims to rationalize the empirical patterns described in Section 2 and provide a framework to quantify the implications of the pending land reform on rural employment in Colombia.

In the first subsection, I present the economy's endowments, the market environment, and the technologies. Then, I define the competitive equilibrium and explain the intuition behind the comparative statics. The last subsection examines the type of land transactions that would occur with sale markets in the long run, as these markets played a crucial role in the change in land concentration that I study in Section 4.

3.1 Environment

Consider a small open economy that produces an agricultural crop q_a and a nonagricultural good q_n. The prices of these two goods are p_a and p_o, respectively, and the non-agricultural good is nontradable. In this economy, individuals consume these two locally produced goods, in addition to a third good that is produced and sold in the city q_c. This city good is a higher-quality substitute for the non-agricultural good that is produced locally.

This local economy is endowed with fixed amounts of labor L, farmland H, and capital equipment K. All individuals are endowed with one unit of labor that is supplied inelastically but have different endowments of land. There are θL agents that have landholdings, and each one of them has a different endowment of hectares given by the distribution $f(h)$. Once landholdings are realized, individuals make their choices in two stages. First, they choose inputs to maximize profits, and then they finance consumption with their disposable income.

In this model, agents are competitive in output markets, and labor is perfectly mobile across sectors. This latter assumption implies a unique equilibrium wage w, which precludes selection as a potential channel for the allocation of employment across sectors. On the other hand, consistent with the patterns documented in Section 2, the capital market features economies of scale. That is, the effective rental price that farmers pay per unit of capital R_i varies across agents and decreases with holding size, i.e, $R_i'(h_i) < 0$. The land market in this model also departs from its competitive benchmark. Notably, land cannot be rented in or rented out. This assumption is consistent with the finding that rental markets in Colombia are thin due to high informality and credit rationing (Gáfaro et al., 2014). Thus, this static model is intended to explain how exogenous shifts in the farm size distribution affect rural employment rather than how this new distribution emerges.[10]

3.1.1 Technologies:

Agriculture: the agricultural good is produced using land h, labor ℓ_a, and capital k_a. The technology features constant returns to scale and is given by

$$q_a = h^{\gamma}[\eta \ell_a^{\rho} + (1-\eta) k_a^{\rho}]^{\frac{1-\gamma}{\rho}}$$

where γ is the output elasticity of land and $\eta \in (0,1)$ captures the relative importance of labor to capital in production. The parameter $\rho < 1$ determines the elasticity of substitution between labor and capital, i.e. $\sigma = \frac{1}{1-\rho}$. As $\rho \to 1$, labor and capital become highly substitutable.

Nonagriculture: the nonagricultural good is produced by a stand-in firm. The technology uses only labor and features constant returns to scale:

$$Q_n = L_n$$

[10]See Section 3.5 for a discussion on the role of sale markets in the consolidation of land in the long run. For a two-sector static model that endogenizes the farm size distribution, see Adamopoulos and Restuccia (2014) and Chen (2017).

where L_n is the aggregate amount of labor units employed in the nonagricultural sector. This characterization reflects the observation that nontradable goods in rural economies are not intensive in capital (Foster and Rosenzweig, 2008). It also precludes the production of value-added products in order to focus on the role of local consumption as a driver of employment in this sector.[11]

3.1.2 Preferences:

I consider an economy where agents consume three types of goods. Two of them are the agricultural $\{c_a\}$ and nonagricultural $\{c_n\}$ goods produced in the rural economy. The third one is a good that is produced and sold in the city $\{c_c\}$, which is a higher-quality substitute for the locally produced (and traded) nonfarm good. Agents choose consumption to maximize

$$\begin{aligned} \max_{c_a,c_n,c_c} \quad & u(c_a,c_n,c_c) = \omega_a log(c_a - \bar{c_a}) + \omega_n log(c_n) + \omega_c log(c_c + \bar{c_c}) \\ \text{s.t.} \quad & p_a c_a + p_n c_n + p_c c_c \leq Y \end{aligned} \tag{1}$$

where ω_j is the relative weight for good j, $\bar{c_a} > 0$ is a subsistence constraint of food consumption, and $\bar{c_c} > 0$.[12] The motivation behind these non-homothetic preferences is to account for the observation that changes in income lead to changes in expenditure shares (see the fourth pattern documented in Section 2). If $\bar{c_a} > 0$, the income elasticity of the agricultural good is less than one. If $\bar{c_c} > 0$, the income elasticity of the city good is greater than one.[13]

3.1.3 The Agent's Problem:

The problem for each agent can be solved in a recursive way. First, she maximizes income. Then, she chooses consumption demands subject to that income. The disposable income of a landless agent is given by the value of their labor endowment $Y = w$. The disposable income of a farmer is given by the sum of the value of their labor and the agricultural profits, $Y = w + \pi_a^*$.

[11] According to the National Household Survey, manufacturing of food, beverages, tobacco, and leather products only contributed with 8% of rural nonfarm employment in 2017. This figure represents a two percentage point increase since 2006.

[12] Intuitively, $\bar{c_c} > 0$ can be interpreted as gifts from relatives who live in the city or public policies that grant access to nonfarm goods from the capital.

[13] Note that these preferences are defined if $Y \geq p_a \bar{c_a}$

Given the market environment, agricultural profits are given by:

$$\max_{\ell_a, k_a} \quad \pi_a = p_a q_a - w\ell_a - R_i(h)k_a \tag{2}$$

To gain insight, I posit that the effective rental price for each farmer i is

$$R_i(h) = r\Big(1 + \frac{1}{h_i^{\nu}}\Big)$$

where r is the benchmark price that is determined in equilibrium and $0 < \nu < 1$.[14]

3.2 Equilibrium

I focus on the competitive equilibrium of the model. Given the market environment, three prices are endogenously determined. These prices are the equilibrium wage w, the benchmark rental price of capital r, and the price of the nontradable good p_n. The price of the agricultural crop p_a is set in the international market, and the price of the city good p_c is set in the city.

Market of nonfarm good: the problem for the stand-in firm that produces the nonfarm good implies that aggregate production is perfectly elastic. This indicates that total production is pinned down by the aggregate consumption of the good,

$$Q_n = C_n = (1-\theta)Lc_n^0(w) + \theta L\left[\int^{h^{-1}(\widetilde{Y})} c_n^1(h)dF(h) + \int_{h^{-1}(\widetilde{Y})} c_n^2(h)dF(h)\right] \tag{3}$$

where $c_n^0(w)$ is the consumption function of landless agents, and $c_n^2(h)$ $(c_n^1(h))$ is the consumption for agents who (do not) purchase the city good.[15] By profit maximization, $\pi_n^* = p_n\ell_n^* - w\ell_n^* = 0$. Thus, since aggregate consumption is positive, in equilibrium, $p_n^* = w^*$.

Labor market: the clearing condition that defines the equilibrium wage in this economy is given by the equality of total labor supply and total labor demand,

$$\bar{L} = L_n + L_a = L_n + \theta L \int \ell_a(h)dF(h) \tag{4}$$

The aggregate labor demand equals the sum of the demand for workers in the non-farm

[14] This functional form implies that $R_i'(h_i) < 0$ and $R_i''(h_i) > 0$. Further, given the parametrization in Section B.1, $R_i \in (r, 2r)$.

[15] $\widetilde{Y}$ is the income threshold at which agents start consuming the city good.

sector L_n and the total workers required by the farmers L_a. Since production in the nonfarm sector is linear in labor, $L_n = Q_n$ as defined in Equation 3.

Capital market: the clearing condition in the capital market is given by,

$$K = K_a = \theta L \int k_a(h) dF(h) \tag{5}$$

where K is the total supply of capital equipment in the economy, and K_a is the aggregate demand for capital across all farmers.

Definition 1. *The equilibrium in this economy is given by input demands $\{L_a, L_n, K\}$, consumption demands $\{C_a, C_n, C_c\}$, and prices $\{w, r, p_a, p_n, p_c\}$; such that:*

1. Input demands in agriculture satisfy the agent's optimization in Equation 2.

2. Consumption demands are consistent with the agent's optimization in Equation 1.

3. Total production of the nonfarm good is given by Equation 3, and $p_n^* = w^*$.

4. Labor market clears following the equality in Equation 4.

5. Capital market clears following the equality in Equation 5.

6. Total farmland equals the sum of endowments across landed agents: $H = \theta L \int h_i dF(h)$.

The exogenous prices p_a^* and p_c^* are set in the international and city markets, respectively.

3.3 Partial Equilibrium Relationships

In what follows, I derive partial equilibrium relationships that are consistent with the empirical patterns presented in Section 2.

Proposition 1. *If $\frac{1}{\sigma} < \gamma$, labor (capital) per hectare decreases (increases) with holding size.*

This proposition indicates that if the elasticity of substitution is high, large producers will substitute labor for capital and become less(more) labor(capital) intensive than their smaller counterparts. This substitution is driven by the nonlinearities in the pricing of capital. If $R_i = r$ for all producers, input intensity would not vary across farm size.

Proposition 2. *Profits per hectare increase with holding size.*

Proposition 3. *There exists an income threshold $\widetilde{Y}$ at which agents start purchasing the city good:*

$$\widetilde{Y} = p_a \bar{c}_a + \frac{p_c \bar{c}_c (1 - \omega_c)}{\omega_c}$$

Further, if $p_a \bar{c}_a < p_c \bar{c}_c$, the expenditure share of the local nonfarm good is hump-shaped.

Given $\bar{c}_c > 0$ and the subsistence constraint $\bar{c}_a > 0$, individuals with low income set $c_c = \bar{c}_c$ and spend their money on food and the locally traded nonfarm good. As income increases, the subsistence constraint ceases to bind, and individuals start substituting expenditure in food with expenditure on the nonfarm good.

At higher levels of income, $Y > \widetilde{Y}$, individuals can afford to purchase the good that is produced and sold in the city ($c_c > \bar{c}_c$). Moreover, given their non-homothetic preferences, they also start spending a higher share of their income on it. If $p_a \bar{c}_a < p_c \bar{c}_c$, this share increases at a higher rate than the rate at which the expenditure share in food decreases. In this case, wealthy individuals finance their consumption of the city good by reducing the expenditure share in both food and the nonfarm good.

3.4 Mechanisms

In this subsection, I describe the main mechanisms behind the effect of an exogenous increase in large-scale concentration on sectoral employment and the equilibrium wage. To illustrate the intuition, consider Panel A of Figure 5, which features the rural labor market. In this figure, the x-axis represents the total supply of workers in the economy, and the red line represents the demand for farm and non-farm labor. The blue line depicts the demand for non-farm labor, which has its origin in the bottom right corner of the figure.

Effect on the equilibrium wage: an increase in large-scale consolidation will have two main effects on the equilibrium wage. First, an increase in the average plot size in the economy will result in a decrease of labor intensity, and thereby a reduction in farm labor demand (i.e. labor intensity effect). Second, consolidation will affect the demand for nonfarm labor since the expenditure in local consumption, $\mathcal{C}_o$, varies with the changes in agricultural profits. In contrast to the labor intensity effect, this shift in the demand for nonfarm labor is ambiguous as it depends on two offsetting effects. On one hand, a larger average plot size increases agricultural profits and thereby the income to spend in the local economy (i.e. profit effect). On the other, the concentration of profits will decrease the income share allocated to local consumption, as wealthier farmers are more likely to replace local consumption with goods that are only sold in the city (i.e. budget share effect). Given these offsetting effects, the shift in the demand for nonfarm labor as a consequence of large-scale consolidation will depend on whether the positive profit effect is larger than the negative budget share effect. If the positive effect is larger, the demand for nonfarm labor will increase (i.e. positive pull response). Conversely, if the budget share effect is larger, there will be a decrease in the nonfarm labor demand (i.e. negative pull response).

To illustrate how these three effects impact the equilibrium wage, refer to Figure 5 . If there is a negative pull response in the nonfarm sector, the demand for nonfarm labor will decrease and there will be an *unambiguous* reduction in the equilibrium wage (Panel B). In contrast, if there is a positive pull response, the net effect on the wage will be *ambiguous* and will depend on whether the increase in the demand for nonfarm labor is large enough to offset the labor intensity effect (Panel C).

Effect on the share of agricultural workers: similar to the analysis on the comparative statics for the wage, the effect of an increase in concentration on the share of agricultural workers will depend on the sign and magnitude of the shift in the demand for nonfarm labor. A positive pull response will result in a decrease in the share of farm workers, as the gap in sectoral wages will promote the reallocation of labor from the farm to the nonfarm sector until a new equilibrium emerges. Meanwhile, a negative pull response will result on an ambiguous overall effect since the type of reallocation will depend on the relative strength of the inward shifts in both curves.

Two important insights emerge from this analysis. First, land consolidation affects the demand for both farm and non-farm labor. Second, the sign of the overall effect on the equilibrium wage is ambiguous. This overall effect depends on the interaction of three main competing effects. The labor intensity effect, the profit effect, and the budget share effect.

3.5 Does Land Get Consolidated Via Sales Markets In The Long Run?

In the model I put forward above, land cannot be rented, and thereby, the size distribution only changes due to exogenous policies. In the long run, however, the size distribution can change through the operation of the sales market, as it was the case studied in Section 4, where land concentration changed due to a trade shock that changed the relative profitability of cash crops. To analyze the role of sales markets in the model proposed above, in this section, I follow Carter and Salgado (1998) and Carter and Zegarra (1994) and compare the competitiveness of landholders in the sales market using a land valuation exercise. The question at hand is: what type of transactions would take place if competitive sales markets are considered?

Given the model in Section 3.1, an agent's willingness to rent land is given by her shadow rental price:

$$\mu_i = \gamma h_i^{(\gamma-1)}[\eta(\ell_i^*)^\rho + (1-\eta)(k_i^*)^\rho]^{\frac{(1-\gamma)}{\rho}}$$

where h_i is the size endowment and ℓ_i^* and k_i^* are the optimal demands for labor and capital, respectively.[16] Thus, we can obtain a measure of land valuation for each agent i as the capitalized stream of the income increments given by this shadow rental price:

$$\Delta_i = \sum_{t=1}^{T} \frac{\mu_{it}}{[1+r]^t}$$

where r is the interest rate and T is the time horizon over which the household expect to receive benefits from that land. This measure represents the net present production value of land for each agent. Notably, if a household wants to buy (sell) one unit of land, Δ_i is the maximum (minimum) price that agent i is willing to pay (accept) for that unit without losing money. The higher this value, the more competitive an agent is in the sales market as land will flow to those agents with a higher willingness to pay.

Proposition 4. *If $\mu_{it} = \mu_i$ and $\gamma < 1$, the net present production value of land increases with plot size*

$$sign\Big(\frac{\partial \Delta_i}{\partial t}\Big) = sign\Big(\frac{\partial \mu_i}{\partial t}\Big) > 0$$

The proposition above implies that large producers place a higher value on land. In particular, given advantages in capital markets, they are more productive and thereby more willing to pay a higher price than small producers for one unit of land. This difference in the valuation of land makes large holders more competitive in sales markets. Thus, since the land will flow from smaller to larger producers, these results suggest that the model would predict a consolidation in large farms if sales markets are considered.

This net present value approach to land valuation is useful and informative about the core income factors that shape the willingness to pay for land. Yet, in comparison to a dynamic general equilibrium model, it abstracts away from risk and intertemporal consumption choices that may also affect the willingness to pay for land (Carter and Kalfayan, 1989; Carter and Salgado, 1998; Carter and Zegarra, 1994). While these two factors may have opposing effects on the competitiveness of smallholders, previous research shows that their consideration may reduce their land valuation even further (Carter and Zim-

[16]Note that if rental markets were thick and considered in the modeled, an agent's willingness to pay to rent land would be equal to the market rental price.

merman, 2000). Thus, the results obtained with a dynamic model will probably enforce the long-run tendency to consolidate land in large farms.

4 The Impact of Large-Farm Consolidation on Rural Employment

4.1 Data Sources and Measurement

To estimate the effect of consolidation on rural labor markets, I gather data from multiple sources and put together a dataset of municipalities for the years 1993 and 2005. Municipalities are the smallest administrative units in the country and are a good approximation to local labor markets since most of their inhabitants work within their geographic boundaries.[17] To focus on rural economies, I exclude large cities and their main agglomerations using the definition proposed by the *Rural Mission* in 2014. This definition classifies municipalities in four categories based on population density and the number of inhabitants living in the town's seat.[18] To account for the segregation and creation of new municipalities across years, I use a consistent unit of observation based on the official boundaries of the municipalities in the year 1993. In my final sample, only 4% of the units of observation are composed of two or more adjacent municipalities.

In what follows, I describe the different sources of information that I use to construct the outcome variables as well as the measure of land concentration. In Section 4.2.1, I provide information on the main characteristics of the municipalities included in my analysis as well as their spatial distribution across the different regions of the country.

4.1.1 Outcomes

To measure impacts on labor market outcomes at the municipality level, I use information from the National Population Census of 1993 and 2005.[19] These two rounds of the census collect information on labor participation and the industry of employment, in addition to the usual demographic and socio-economic characteristics collected in the other rounds.[20] For each municipality, I compute the unemployment rate as the ratio of the number of unemployed individuals to the total population that is economically active. Similarly, I

[17] According to the population census of 2005, almost all (95%) the workers who lived in rural municipalities worked in the same municipality where they lived.

[18] See Ocampo (2014) for more details on this classification.

[19] The National Department of Statistics granted me access to the microdata of these two rounds of the census per confidential agreement. I had access to this microdata at the data center of the Department of Statistics during the second semester of 2019.

[20] These two rounds of the Population Census have been shown to be comparable despite their differences in implementation (Le Roux, 2013; Mallarino, 2007; Jaramillo and Ibáñez, 2005). The most recent census, conducted in 2018, did not collect information on sectoral employment.

construct the share of agricultural workers using information on the economic sector of the main job in the last 15 days.[21] In order to classify employment by sector, I define agricultural jobs as those activities that make an intensive use of land to produce crops and raw materials, raise livestock and poultry, or farm fish. Any other activities, including the processing of food and beverages, are classified as non-agricultural jobs.

Although the data on industry refers to the last 15 days prior to the census, it is unlikely that the differences in the share of agricultural workers across locations are largely driven by the timing of data collection. For one, the roll-out of the population census across regions is not correlated with any specific raining season. Moreover, as shown in Figure A1 in the appendix, for most regions of the country, the share of agricultural workers in rural economies does not vary much throughout the year. This relative stability of sectoral employment is driven by the existence of two raining seasons during the year as well as the existence of thermal floors that guarantee that each region can produce a wide variety of crops with different time-to-maturity and harvest cycles.[22]

To conduct the analysis on wages, I use individual-level data from the National Household Surveys (ENH) of 1998 and 2009. These surveys are representative at the national and department level and include a large set of rural municipalities that are selected at random each round. The employment module of this survey collects cross-sectional data for a random sample of workers on features such as occupation, hours worked, and income. Using this information, I compute the hourly wage for each worker as the ratio of her monthly wage to the number of hours worked during the last month, and convert nominal to real values using the price index of 1993.

To compute these three outcomes, I focus on individuals between 15 and 65 years old.

4.1.2 Land Concentration

To construct measures of land concentration, I use data from the national cadastre system for the years 1993 and 2005. This system is a census of all the properties in the country with detailed information on the location of the plot, the type of holder, and the plot's size.[23] This national cadastre is managed by five different agencies that are in charge with

[21] Data on sectoral employment for 2005 was collected for a random sample of workers, which is representative within narrow cells defined by the municipality and the area of residence (i.e. county-seat vs disperse area) (DANE, 2005). As not all workers provided information on industry, I calculate employment shares by sector as the ratio of the number of workers employed in each sector to the total number of workers with information on industry. In the appendix, I present evidence that suggests that non-response rate is not a threat to the internal validity of my estimates.

[22] This observation also holds in the economies of the Central region, where agricultural labor is expected to be more dependent on the seasonality of coffee production.

[23] The data for 1993 was generously provided by Fabio Sanchez from the Department of Economics at the University of the Andes in Bogota, Colombia. I obtained the data for 2005 with the purchase of the report

the creation and updating of the registry for different parts of the country. In my analysis, I use the information managed by the department of Antioquia and the National Institute of Geographic Information (IGAC), which constitute the whole universe of rural municipalities with cadastral data.[24] For each municipality and year, I have information on the number of properties and their total area across 13 different plot's size ranges. For instance, the number of plots that are between 3 and 5 hectares and their corresponding surface.[25] To focus on privately held land, I exclude the records of indigenous reserves as well as those properties that belong to the State.[26]

Using this cadastral data, I construct a measure of land concentration on the right tail of the plot size distribution (i.e. large scale). Specifically, for each municipality and year, I calculate the share of area in large plots. In order to define what a 'large' plot is for each municipality, I use as reference the town's average family farm unit, known as *UAF* for its acronym in Spanish. This unit is a policy instrument that represents the minimum plot size that a family needs to remunerate the factors of production and generate an income surplus, given the agro-ecological conditions of the plot's location.[27] Following Machado and Suarez (1999), I define a "large" plot as one that has a size of at least 10 family farm units. Specifically, based on the 13 size ranges that I observe, I construct the share of area in large plots for each municipality i as:

$$\omega_i^{large} = \omega_i^a \quad if \quad \kappa_i \in [a, b) \tag{6}$$

where κ_i refers to 10 times the family farm unit, (a, b) refer to the lower and upper bounds of the observed size ranges, and ω_i^a refers to the share of area in plots with a size of at least

conducted by IGAC, entitled 'Atlas de la Distribución de La Propiedad Rural en Colombia'.

[24]The other three agencies are responsible for the cadastral data of the largest capital cities in the country: Bogota, Medellin, and Cali (IGAC, 2012).

[25]The size ranges for which this information is available are: less than 1 hectare, 1 to 3 hectares, 3 to 5 hectares, 5 to 10 hectares, 10 to 15 hectares, 15 to 20 hectares, 20 to 50 hectares, 50 to 100 hectares, 100 to 200 hectares, 200 to 500 hectares, 500 to 1000 hectares, 1000 to 2000 hectares, and more than 2000 hectares.

[26]The concept of land concentration is inappropriate for the case of indigenous reserves since this land is held under collective tenure. In contrast to the data for 2005, the data for 1993 does not provide detailed information to determine whether it also excludes the records of Afro-Colombian communities and National Parks. Therefore, in a future version of this draft, I will check if the results of my analysis are robust to different filters that I apply to the data.

[27]The family farm unit was initially created by the Law 135 of 1961 with the intent to guide the allocation of public vacant lands. It represents the minimum plot size that is required, in a given region, to produce an income of three monthly minimum wages and a disposable income after paying land rent payments of at least 66%. This unit takes on different values depending on the type of agro-ecological zone and land use, so the average family farm unit for each municipality is a weighted average of the different values across production systems and zones (Departamendo Nacional de Planeación, 2000). For details on the distribution of the average family farm unit across the municipalities in my study population, see Section 4.2.1.

a hectares.[28]

In addition to this measure of *large-scale* concentration, I also use this data to calculate the Gini index of landholdings, which is a measure of *overall* concentration across the whole plot distribution. This index, calculated with binned data, is a lower bound for the index calculated with the microdata of plots, given that concentration is higher within size ranges than across them.[29]

One of the advantages of using the cadastral data to construct measures of land concentration is that the system records information based on possession as opposed to ownership. That is, all the plots held by individuals and firms are classified as private regardless of the existence of formal titling. This feature of the cadastre is crucial to guarantee that my data represents the real concentration of land in the country, given that an important fraction of the farmers have informal titles and are settled in public vacant lots (**?**). Additionally, since the data gathered upon *formation* of the cadastre is collected and updated by personnel in the field (IGAC, 1988), the plot size measures recorded in the cadastre are less likely to be afflicted by the self-reporting bias that is common in survey data.[30]

4.1.3 Municipal Characteristics

In addition to the data on outcome and land consolidation, I compile a set of municipality characteristics using different data sources. I calculate the municipality's average slope and elevation using the Data Elevation Model from NASA together with GIS software and shapefiles with 1993 boundaries.[31] Similarly, I obtain a measure of the most prevalent climatic zone in the municipality using the climatic maps provided by the National Geographic Institute (IGAC). To define what a large plot means for each municipality given terrain and market conditions, I use as reference the average family farm unit calculated by the National Department of Statistics (DANE). I also match each municipality with its respective *province* and *department* using the official administrative and political divisions.

In order to account for the prevalence of forced displacement and the increase in con-

[28]For instance, if the average family farm unit is 32 hectares, the share of area in large plots equals the share of area in plots of at least 20 hectares. Since this measure based on size ranges is less precise than the one I would obtain with microdata, in a future version of this draft, I will check whether the main results of the paper are robust to different versions of Equation 6.

[29]Specifically, using data for 2005, I note that the average Gini index across the municipalities in my sample is 0.69 if I use the microdata and 0.37 if I use the binned data across size ranges.

[30]One important caveat of this data, however, is that resource limitations prevent the cadastral agencies from carrying out the updates during the established 5-year window for all the municipalities. In fact, Pinzón and Fonti (2007) find that only 31% of the municipalities were fully up-to-date in 2005. Yet, as I show in the appendix in general, the probability of having an updated cadastre is uncorrelated with the instrumental variable that I use in my empirical analysis .

[31]The Data Elevation Model from NASA has a pixel resolution of 1 arc-second (approximately 30 meters).

flict intensity during the period of study, I use the administrative data compiled by the National Memory Center to calculate the cumulative number of conflict-related events, the cumulative number of murder victims, and the share of population that was forcefully displaced from each municipality between 1993 and 2005.[32] As forced displacement changes the total population in the municipality, I use as denominator the total population in 1993. Finally, to measure initial levels of urbanization, I digitize population indicators of the 1985 census. Specifically, the total number of inhabitants and the share of rural population in each town.[33]

4.2 Empirical Analysis

In this section, I outline the research design that I use to estimate the effects of land consolidation on the structural transformation of rural Colombia as well as other labor outcomes such as unemployment rates and wages. I first describe the study population and then introduce the identification strategy and discuss its main assumptions.

4.2.1 Study Population

The study population in this analysis consists of 590 municipalities across 21 *departments* of the country. These 21 departments correspond to 50% of the country's area and 95% of the national population.[34] Once I exclude large cities, my study sample is composed of 73% of the rural economies in these departments.[35] These towns are spread in similar proportions across lowlands, hills and highlands, and hence are representative of a large set of agricultural goods produced in the country (see Figure 6).

In Table 1, I present summary statistics that describe the main characteristics of the sample before the observed changes in land concentration (i.e. 1993). These municipalities were located about 79 kilometers from the department's capital and had a mean area

[32]Conflict-related events include the total number of attacks, incursions, and assaults on the civilian population perpetrated by illegal armed groups. For more details on this data, see Centro Nacional de Memoria Historica (2013).

[33]These are the only indicators that are reported in the Census Report of 1985. Due to travel restrictions during the pandemic, I was unable to access the microdata to calculate additional indicators, including predetermined levels of the outcomes of interest.

[34]I exclude the municipalities from the departments of Amazonas, Arauca, Casanare, Caquetá, Chocó, Guainía, Guaviare, Putumayo, San Andrés, Vaupés, and Vichada for at least one of the following reasons. First, the department was not part of the agricultural frontier in the 1990s. Second, the rural lands in the department consisted mostly of indigenous reservations and afro-colombian lands, where rights are communal and hence the concept of consolidation is not appropriate. Finally, cadastral data for these places is either non-existent or unreliable.

[35]The remaining 27% of rural economies in these departments did not have data on land concentration or sectoral employment for both years of analysis.

of 537 square kilometers. In comparison to large cities, these municipalities had a low population count with an average of 15,543 inhabitants. A large share of these individuals (69%) lived in disperse areas outside the county-seat, and the majority of workers (67%) worked in agricultural activities. Most of these rural towns had a low unemployment rate (2.5%) and a wide spread in the wage perceived by workers. While the average hourly wage (COL$524) was slightly above the minimum legal wage at the time, about two-thirds of the workers were making less than that.[36]

Table 1 also shows that concentration at the top of the plot size distribution was already prevalent in 1993. Specifically, only a small minority of towns (6%) had zero hectares in large farms, while 75% of towns had at least 11% of their area in large farms. As explained in Section 4.1.2, I use the family farm unit of each municipality as reference for the threshold that defines a large plot size. In my sample, this family farm unit ranges from 2.5 to 126 hectares and has a mean of 22 hectares. Finally, between 1993 and 2005, all the municipalities displayed a general change in the concentration of land. Overall concentration, measured by the gini index, increased in 40% of the municipalities, and concentration in large plots increased in 18% of the towns.

4.2.2 Identification Strategy

From an evaluation standpoint, a major challenge to assess the effect of a change in land concentration on rural labor markets consists on the multiple sources of endogeneity that relate consolidation with my outcomes of interest. For one, large-scale policies of land allocation are usually targeted and heavily influenced by political groups (Albertus, 2015; Faguet et al., 2020; Machado, 2013). More generally, land concentration is an equilibrium outcome that not only depends on factors that affect labor markets (i.e. initial settlement patterns, distance to markets), but is also driven by employment outcomes themselves. For instance, consolidation could be the result of smallholders exiting agriculture in response to booming opportunities in the non-farm sector. One additional endogeneity threat that is specific to the Colombian context pertains to the intensification of armed conflict during the period of analysis. This upsurge in violence, driven by the expansion of paramilitary groups throughout the country, affected both rural labor markets and the concentration of land. On the one hand, conflict caused a massive displacement of the rural workers to the cities, and induced changes in agricultural production and labor supply (Arias et al., 2018; Fernández et al., 2014).[37] On the other, a large proportion of the

[36] This figure corresponds to a wage of COL$3,806 per hour (U$1 per hour) in 2019 pesos (dollars).

[37] Official records from the National Center for Victims state that more than 4.3 million individuals were displaced between 1993 and 2005. A large majority of these refugees were important leaders of the commu-

abandoned lands were illegally appropriated by armed groups who retained them or sold them to other people (Ibáñez and Muñoz-Mora, 2010; Centro Nacional de Memoria Historica, 2012).[38] In line with historical accounts, nearly all the municipalities in my sample (99.8%) experienced at least one conflict-related event between 1993 and 2005. All these municipalities registered forced displacement and had, in average, 19% of their population displaced (see Table 2).

In addition to endogeneity, a second identification concern refers to the potential existence of spillover effects due to market integration across regions. For instance, if consolidation encourages a *large* number of workers to move to control economies in search of employment, the equilibrium wage in those locations may change even if land concentration stayed the same.[39] Yet, while migration is a common coping mechanism in Colombia (Arteaga and Ibáñez, 2018), two main facts suggest that spatial spillovers across rural economies were weak during this period. First, most of the rural migrants were refugees who moved to large cities in search of assistance and anonymity (de Planeación, 2008). Further, since rural markets were poorly integrated (Nupia, 1997), it was rare for workers to commute to other economies for work. In fact, using the population census of 2005 I find that most of the rural workers (95%) worked in the same municipality where they lived. Hence, since spatial spillovers do not seem to be a major threat to the internal validity in this setting, in the rest of this section I focus on addressing endogeneity concerns.

To make progress on the identification of causal estimates, I use an instrumental variable approach. In particular, using panel data, I exploit the differential exposure of rural economies to a demand shock that rendered a land-use change towards land-intensive crops. This change in demand was driven by a process of trade liberalization in the 1990s that improved the access of Colombian products to international markets. Specifically, during this decade the country completed the first two Free Trade Zones negotiations and promoted the production and exports of cash crops through the use of subsidies and tax-exemptions.[40] Since this process of market integration changed the profitability of

nity, who worked in different sectors of the economy such as agriculture, education, and commerce (Ibáñez, 2008).

[38] According to the Official Group of Historical Memory, 8.4 million hectares were illegally seized between 1994 and 2010 (Centro Nacional de Memoria Historica, 2012). This area corresponds to 32% of the total land with agricultural vocation in the country.

[39] The existence of spatial spillovers violate the non-interference condition known in the treatment effect literature as the Stable Unit Treatment Assumption; i.e. *SUTVA*, (Rubin, 1978). In this example, if migration affects rural labor markets in control economies, the average treatment effect of consolidation cannot be identified with the difference in mean outcomes, even if consolidation occurred as part of a random experiment in the first place.

[40] See Jaramillo (2002) for more information on the agricultural commercial policy during the 1990s. Some of the policy measures that were implemented to promote agricultural exports were the Law 693 of 2001, the Law 818 of 2003, and the *price stabilization funds*.

crops, during these years the agricultural sector experienced a productive transformation (Jaramillo, 2002). On one hand, there was an increase in the area used to produce perennials such as bananas, flowers, sugar cane, and oil palm.[41] On the other, there was an increase in the area used to raise beef cattle, especially in those regions that were not as suitable for the production of lucrative export crops (Balcázar, 2003).[42] These crops are usually land-intensive due to a high degree of vertical integration, or, in the case of cattle, a maximum carrying capacity per hectare.[43] Thus, as the production of these various crops is feasible in multiple regions of the country, this productive transformation brought with it a general change in the scale of agrarian operation.[44]

To measure the differential exposure of each geographic area to this demand shock, I use topographic features of the economy that are related to the financial viability of production at scale. In particular, I use a precise measure of the municipality's average slope since, for any given crop, it is cheaper to build and maintain infrastructure for large-scale production in even terrains.[45] Likewise, flat terrains are more attractive for mechanization and extensive cattle ranching, as the carrying capacity per hectare decreases with the field's gradient.[46] Of course, since topography also influences crop suitability, the municipality's slope is certainly correlated with initial levels of development that affect labor markets. Yet, since my empirical strategy consists of comparing the *change* in outcomes before and after the demand shock and across economies with distinct exposure, these pre-existent differences are not necessarily a threat to the validity of the instrument.

[41]See Figure 7 for the trends in harvested area for some of these crops.

[42]In addition to the commercial agreements, in 1997 the government created a National Program to overcome sanitary barriers to export cows. Hence, the national increase in the area used for this activity was driven in large by the expected benefits of entry to international markets in the near future. Additionally, cattle ranching became an alternative in those places where cash crops were not suitable (Balcázar, 2003), since it is an attractive investment that allows farmers to retain the property rights of their land while incurring in low input and tax cost (Kalmanovitz, 2015, 2002).

[43]Since cattle has a maximum carrying capacity per hectare, its production is only profitable after a minimum plot size threshold.

[44]The production of these diverse crops is feasible in multiple regions of the country. For instance, flowers are produced in the highlands, oil palm is produced in the lowlands, and beef cattle can be produced in the lowlands, the inter-andean valleys, and the slopes (Federación Colombiana de Ganaderos, 2018). During this period, the economies of the slopes also experienced changes in the scale of agrarian operation in response to the decrease in the international price of coffee. Specifically, there was a change in land use that resulted in the fragmentation of coffee farms (Balcázar, 2003; García, 2003).

[45]To construct the municipality's average slope, I overlay the shapefile of municipalities with the raster of the Data Elevation Model from NASA, which has a high resolution of 1 arc-second per cell (30 meters approximately). Each municipality's average slope is given by the (area) weighted average of the slope across all the raster cells that intersect with its surface.

[46]Some examples of the infrastructure that is required to produce export crops such as oil palm and flowers are irrigation systems and greenhouses. According to qualitative interviews conducted on various regions of Colombia during 2019, the carrying capacity of cattle in flat lands is about 3 cows per hectare compared to 2 (or less) cows per hectare in the slopes.

To illustrate this, consider the following second and first-stage equations where D refers to land concentration, and Y refers to any outcome of interest for municipality m in province p at time t:[47]

$$Y_{mpt} = a_m + \beta \hat{D}_{mpt} + \delta_{p,t} + u_{mpt} \quad ; \quad D_{mpt} = a_m + \gamma(S_{mp} \times T_t) + \delta_{p,t} + \eta_{mpt} \tag{7}$$

In general, the instrument is given by the interaction of the demand shock, driven by market integration, and the location's differential exposure to this change.[48] However, since for $t = 2$ the shock is perfectly collinear with the time trend (T_t), all the variation in the instrument is given by its cross-sectional component; i.e. the municipality's average slope (S_{mp}).[49] By including municipality fixed effects (a_m), I control for differences in initial outcome levels correlated with topography through other channels besides the financial viability of production at scale (i.e. crop choice). Consequently, the exclusion restriction, required for identification, is that the municipality's slope does not lead to differential changes in local labor markets.

While this empirical strategy addresses issues driven by simultaneity and time-invariant confounders, there are still a few endogeneity concerns that remain. In particular, given the close relationship between topography and crop suitability, one concern consists of the crop-specific income effects caused by the trade liberalization. For one, the changes in land use during this period were driven by changes in the profitability of export crops. Moreover, the government implemented sector-specific subsidies to alleviate the burden of liberalization on producers of importable crops.[50] To address this concern, I control for time-trends that are common across municipalities with similar agro-climatic conditions. Specifically, I use *province-year* fixed effects $(\delta_{p,t})$.[51] These *department* subdivisions are composed of an average of 8 contiguous municipalities and have been used historically to plan and develop environmental and territorial policies (DANE, 2014). Hence, to the extent that economies within a province faced the same social and economic shocks, the inclusion of these trends accounts for the income effects.

In regards to the increase in conflict intensity, one potential concern is that the upsurge was more severe in the mountains, since, for decades, they have provided a natural shelter for illegal groups (Centro Nacional de Memoria Historica, 2013). In fact, when

[47] In line with the applied literature, this outcome equation assumes that the relation between concentration and outcomes is linear in parameters and that both the time-invariant (a_m) and time-variant unobservables (u_{mpt}) enter the outcome equation in additive form.

[48] This instrument is similar in spirit to the Bartik instrument that leverages the differential exposure of subnational economies to common nation-wide industry shocks (Goldsmith-Pinkham et al., 2020).

[49] In this setting, the years 1993 and 2005 refer to the pre-shock and post-shock periods, respectively.

[50] See, for instance, the Andean Price Band System enacted in 1993.

[51] There are 100 *provinces* across the 21 *departments* that are part of my study population.

I estimate correlations between the municipality's slope and the change in conflict, I find that is the case. Locations with steeper slope were more likely to have an increase in homicides, displacement, and attacks (see Table A1 in the appendix). Yet, consistent with the fact that municipalities within a province share similar trends, most of these correlations vanish once I control for province-year fixed effects. The only conflict-related event that remains significant is the cumulative share of forcefully displaced population during this period. Hence, to account for this potential confounder, I include this measure as a covariate (C_{mpt}) and estimate:

$$Y_{mpt} = a_m + \beta \hat{D}_{mpt} + \delta_{p,t} + \alpha_1 C_{mpt} + u_{mpt} \quad ; \quad D_{mpt} = a_m + \gamma(S_{mp} \times T_t) + \delta_{p,t} + \alpha_2 C_{mpt} + \eta_{mpt} \tag{8}$$

To give an idea of the variation I use to estimate β, in Figure 8 I show the spatial distribution of the instrument and its values after partialling out the fixed effects and covariates. In all estimations, standard errors are clustered at the municipality level to allow for serial correlation. The main identifying assumption in Equation 8 is that, once I adjust for differences in conflict intensity, topography does not lead to different labor market trends across economies of the *same* province.[52] While this assumption is untestable, in the next section I provide reassuring evidence that suggests that economies of the same province had similar population pre-trends, regardless of their topography.

Finally, when it comes to the analysis on wages, I use a repeated cross-section of individuals and estimate the following modified version of Equation 8, where Y_{imjt} refers to the log hourly wage of worker i in municipality m, department j, at time t:

$$Y_{imjt} = a_m + \beta \hat{D}_{mjt} + \delta_{j,t} + \alpha_1 C_{mjt} + u_{imjt} \quad ; \quad D_{mjt} = a_m + \gamma(S_{mj} \times T_t) + \delta_{j,t} + \alpha_2 C_{mjt} + \eta_{mjt} \tag{9}$$

This repeated cross-section consists of a random sample of workers from 49 municipalities that were included in the survey both years. While these 49 municipalities are representative of my initial study population, only a few of them are contiguous municipalities of the same province. Hence, I use department-year fixed effects ($\delta_{j,t}$) to account for crop-specific income effects, and control for the total and share of rural population in 1985 to allow for differential trends for municipalities with different initial urbanization rates. These last two control variables are important in this analysis on wages because, as I show in Table 7, the town's slope led to differential trends in population growth across municipalities of the same *department* even before the trade shock took place.

[52] Formally, $E(s_{mp}\Delta u_{mp}|\Delta C_{mp}, \delta_p) = 0$. This assumption is weaker than the one required for the within estimator, which states that the *idiosyncratic* component of land consolidation across municipalities of the same province is not correlated with labor market shocks; i.e. $E(\Delta D_{mp}\Delta u_{mp}|\Delta C_{mp}, \delta_p) = 0$.

In contrast to the analysis in Equations 8, I use a two-sample 2sls approach to estimate the Equations in 9. This approach allows me to improve the precision of my estimates by increasing the sample of municipalities in the *first-stage*. That is, I use the subsample of 49 municipalities with data on workers's wage in the *second-stage* and the remaining municipalities in the *first-stage*.[53] The consistency of this procedure requires the equality of the first-stage coefficient across both samples, which I document in the results section. To obtain standard errors that are robust to heteroskedasticity, I follow the procedure proposed by Pacini and Windmeijer (2016).

If the identifying assumptions for Equations 8 and 9 hold, β identifies the average marginal effect on the subpopulation of economies that respond to encouragement from the instrument (i.e. compliers).[54] This estimand is different from the average marginal effect on the whole set of economies in my study population. But, it is still an important object of interest that informs policy in those occasions where land consolidation is driven by the physical and financial viability of production at scale.

4.3 Results

Table 3 documents the average effect of a change in the concentration of land in large farms on the structure of rural employment and unemployment rates. For each outcome, I present the 2sls estimates of different specifications that include different sets of covariates. The first stage results confirm that the instrument is strong enough to guarantee that the bias of the 2sls estimates is at most 10% of the bias of the OLS estimates. Specifically, the Kleibergen-Paap F-statistic, which tests for instrument relevance in the presence of correlated errors, is always above the rule-of-thumb value of 10. Yet, since surpassing this threshold is not enough to guarantee correct inference (i.e. size control) (Andrews et al., 2019; Stock and Yogo, 2005) , I also report inference results that are robust to weak-instruments. Specifically, the p-value and 95% confidence set of the Anderson-Rubin test.

The results indicate that the consolidation of land in large farms decreased the share of agricultural workers in rural economies and increased the unemployment rate. In particular, an increase in consolidation of one standard deviation resulted in a 10 percentage point decrease in the proportion of agricultural workers and a 14 percentage point increase

[53] The two-sample 2sls approach was first proposed by Angrist and Krueger (1995) as a procedure to obtain 2sls estimates using two independent samples, one for each stage. For more information, see Inoue and Solon (2010), Pacini and Windmeijer (2016), Choi et al. (2018), and Angrist and Krueger (1992).

[54] To be precise, monotonicity is an additional assumption that is required to interpret β as the local average marginal effect of land consolidation on the outcome of interest. This assumption, untestable in nature, implies that while increasing the average slope can either encourage land consolidation or have no effect at all, it cannot discourage consolidation relative to lower slope values (Kennedy et al., 2019).

in the unemployment rate. These effects are significant at the 95% confidence level and remain relatively unchanged when controlling for the share of forcefully displaced population. These results are also robust to inference methods that allow for weak-instruments. Notably, the p-value of the Anderson-Rubin test always rejects the null hypothesis that land concentration has no impact on the outcome of interest. In Table 4, I show that these results are robust to different types of standard errors, including one that allows for both the correlation of the error term for the same municipality across time and the correlation of the error term across municipalities of the same department at the same point in time (i.e. two-way clustered by municipality and department-year).[55]

My empirical strategy relies on the assumption that, conditional on the covariates, the town's slope does not drive differential trends in labor markets across municipalities of the same province. Thus, while this assumption cannot be tested, using the available data, I conduct a couple of exercises that provide some reassurance that this is a reasonable assumption to make. First, I show that the province-year fixed effects play an important role in controlling for the expected bias generated by the crop-specific income effects *during* my period of study. In particular, since the crops that received higher subsidies were those that were produced in the lowlands (Jaramillo, 2002), comparing towns in the lowlands with towns in the highlands without accounting for such income trends underestimate the negative impact of land concentration on unemployment rates (see Table A2 in the appendix). Second, I test whether the town's slope led to differential trends *before* my period of study, and find that, once I control for conflict events, towns of the same province did not display significant pre-trends in variables such as population size and the share of rural population (see Table 5), which are measures of urbanization that relate to my outcomes of interest.[56]

In Table 6, I present the average effect of the consolidation of land in large farms on wages. This analysis is conducted using repeated cross-sections of the National Household Survey and hence the unit of analysis is the individual in each municipality. For each specification, I present evidence of the equality of the first-stage coefficient across both samples, which is an additional assumption for the consistency of the two-sample 2sls estimates. Similar to Table 3, I also present the Kleibergen-Paap F-statistic and the inference results that provide correct coverage with potentially weak instruments.[57] The standard

[55]For details on multiway clustering, see Cameron et al. (2011).

[56]Currently, I do not have access to the microdata of the 1985 Census to investigate if there are pre-existing trends in my outcomes of interest. To conduct this analysis, I need to access the data at the National Department of Statistics in Bogota, Colombia. Due to travel restrictions driven by the pandemics, this exercise is not yet feasible. However, I plan to include it in a future version of this draft.

[57]To obtain these inference results for potentially weak instruments, I use the Stata command *weaktsiv*, which is based on the procedure proposed by Choi et al. (2018).

errors reported in Table 6, however, are only robust to heteroskedasticity, since, to date, there are no inference procedures for two-sample 2sls that allow for intra-cluster correlation in the presence of weak instruments. The results indicate that the consolidation of land in large farms had no significant effect on the wage of rural workers. The coefficient of interest is negative, but remains insignificant across different specifications and the use of inference methods that allow for weak-instruments.

5 Conclusions and Discussion

This paper provides theoretical and empirical evidence of the effect of concentration of land in large farms on rural labor markets in Colombia. The main idea behind this analysis is that large-scale concentration can affect the equilibrium wage through four main empirical facts that relate the scale of agrarian operation with the demand for farm and nonfarm workers. To analyze the interaction of these demand mechanisms, I propose a general equilibrium model of rural economies with two sectors. This model highlights the importance of local multiplier effects on the creation of employment, as two-thirds of rural nonfarm workers work in non-tradable industries. Using this framework, I show that the overall effect on the structure of rural employment and the wage depends on the interaction of three competing effects: a labor intensity effect, a profit effect, and a budget share effect. And that, in general, large-scale concentration can reduce worker's welfare if it is not accompanied by an increase in the demand for nonfarm labor that is large enough to offset the labor intensity effect.

To estimate the overall impact of large-scale concentration on rural labor markets in Colombia, I use an instrumental variable approach. In particular, I exploit the differential exposure of municipalities to a common trade shock in the 1990s that increased the area used to produce land-intensive crops. The results suggest that the concentration of land in large farms had significant and sizeable effects on rural labor markets. Specifically, it prompted the movement of labor out of agriculture and increased unemployment rate.

While this paper is a step forward toward understanding the effect of large-farm concentration on rural labor markets, it opens several questions for future research. For example, what is the contribution of each of the mechanisms? And how heterogeneous are the effects by type of crop? The theoretical framework that I propose can also be used to quantify the effect of the upcoming land reform on welfare and the structural transformation of rural Colombia. This is an extension I am currently working on.

References

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *The American Economic Review*, 2014, *104* (6), 1667–1697.

_ **and** _ , "Land Reform and Productivity: A Quantitative Analysis with Micro Data," *American Economic Journal: Macroeconomics*, 2020, *12*, 1–39.

_ **, Loren Brandt, Jessica Leight, and Diego Restuccia**, "Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China †," 2019.

Adams, Dale, "Colombia's Land Tenure System: Antecedents and Problems," *Land Economics*, 1966, *42* (1), 43–52.

Albertus, Michael, *Autocracy and Redistribution: The Politics of Land Reform*, Cambridge: Cambridge University Press, 2015.

Ali, Daniel, Klaus Deininger, and Anthony Harris, "Does Large Farm Establishment Create Benefits for Neighboring Smallholders? Evidence from Ethiopia," *Land Economics*, 2019, *95* (1), 71–90.

Andrews, Isaiah, James H Stock, and Liyang Sun, "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, 2019, *11.*

Angrist, Joshua D. and Alan B. Krueger, "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 1992, *87*, 328–336. Goal: this is the paper that develops the two-sample IV estimator.

_ **and** _ , "Split-sample instrumental variables estimates off the return to schooling," *Journal of Business and Economic Statistics*, 1995, *13*, 225–235. Goal: this paper derives assumptions and properties of the split-sample IV estimator. I also has some applications.

Arias, María Alejandra, Ana María Ibáñez, and Andrés Zambrano, "Agricultural production amid conflict: Separating the effects of conflict into shocks and uncertainty," *World Development*, 2018.

Arteaga, Julián and Ana María Ibáñez, "Migración en La ELCA: ¿Quiénes Migran, Por Qué Motivos y Cuáles Son Los Beneficios Potenciales?," Documentos CEDE 015990, Universidad de los Andes - CEDE January 2018.

— **, Carolina Castro Osorio, Diana Cuéllar, Ana María Ibañez, Rocío Londoño Botero, Manuel Murcia, Javier Neva, Ángela Nieto, Dora Inés Rey, and Fabio Sánchez**, "Fondo de Tierras del Acuerdo Agrario de la Habana: Estimaciones y Propuestas Alternativas," Documentos CEDE 015630, Universidad de los Andes - CEDE June 2017.

Avila, Antonio Flavio Dias and Robert E. Evenson, "Chapter 72 Total Factor Productivity Growth in Agriculture. The Role of Technological Capital," 2010.

Balcázar, Álvaro, "Transformaciones en la agricultura colombiana entre 1990 y 2002," *Revista de Economía Institucional*, 2003, *5* (9), 128–145.

Bazzi, Samuel, "Wealth Heterogeneity and the Income Elasticity of Migration," *American Economic Journal: Applied Economics*, 2017, *9*, 219–255.

Bottazzi, Patrick, David Crespo, Leonard Omar Bangura, and Stephan Rist, "Evaluating the livelihood impacts of a large-scale agricultural investment: Lessons from the case of a biofuel production company in northern Sierra Leone," *Land Use Policy*, 2018, *73* (C), 128–137.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, 2016, *106*, 1320–1365.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Robust Inference With Multiway Clustering," *Journal of Business Economic Statistics*, 2011, *29* (2), 238–249.

Carter, Michael, "Identification of the Inverse Relationship between Farm Size and Productivity : An Empirical Analysis of Peasant Agricultural Production," *Oxford Economic Papers*, 1984, *36* (1), 131–145.

— **and John Kalfayan**, "A General Equilibrium Exploration of The Agrarian Question," Technical Report 1989.

Carter, Michael R and Eduardo Zegarra, "Land Markets and the Persistence of Rural Poverty : Post-Liberalization Policy Options," 1994, *8*, 125–152.

Carter, Michael R. and Frederick J. Zimmerman, "The Dynamic Cost and Persistence of Asset Inequality in An Agrarian Economy," *Journal of Development Economics*, 2000, *63* (2), 265–302.

Carter, Michael R and Ramón Salgado, "Land Market Liberalization and The Agrarian Question in Latin America," 1998.

Caselli, Francesco and Wilbur John Coleman II, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 2001, *109* (3), 584–616.

Centro Nacional de Memoria Historica, *Justicia y paz. Tierras y territorios en las versiones de los paramilitares* 2012.

__ , *¡Basta Ya! Colombia: Memorias de Guerra y Dignidad* 2013.

Chavas, Jean-Paul, "Chapter 5 Structural change in agricultural production: Economics, technology and policy," in "Handbook of Agricultural Economics," Vol. 1 of *Agricultural Production*, Elsevier, 2001, pp. 263–285.

Chen, Chaoran, "Untitled land, occupational choice, and agricultural productivity," *American Economic Journal: Macroeconomics*, 2017, *9*, 91–121.

Choi, Jaerim, Jiaying Gu, and Shu Shen, "Weak-instrument robust inference for two-sample instrumental variables regression," *Journal of Applied Econometrics*, 2018, *33* (1), 109–125.

Collier, Paul and Stefan Dercon, "African Agriculture in 50 Years: Smallholders in a Rapidly Changing World?," *World Development*, 2014, *63*, 92–101.

DANE, "Propuesta de Codificacion de Nuevas Divisiones Administrativas," Technical Report 2014.

Davis, Kyle F., Paolo D'Odorico, and Maria Cristina Rulli, "Land Grabbing: A Preliminary Quantification of Economic Impacts on Rural Livelihoods," *Population and Environment*, 2014, *36* (2), 180–192.

de Estadistica DANE, Departamento Administrativo Nacional, "Manual Tecnico 2005," Technical Report 2005.

De Janvry, Alain and Elisabeth Sadoulet, "Rural Development in Latin America: Relinking Poverty Reduction to Growth," in Michael Lipton and Jacques Van Der Gaag, eds., *Including The Poor*, The World Bank, 1993, chapter 11.

de Planeación, Departamento Nacional, "Evaluación Familias en Acción para Población en Situación de Desplazamiento," Technical Report 2008.

Deininger, Klaus, "Challenges posed by the new wave of farmland investment," *The Journal of Peasant Studies*, 2011, *38* (2), 217–247.

_ **and Fang Xia**, "Quantifying Spillover Effects from Large Land-based Investment: The Case of Mozambique," *World Development*, 2016, *87* (C), 227–241.

_ **, Derek Byerlee, Jonathan Lindsay, Andrew Norton, Harris Selod, and Mercedes Stickler**, "Rising Global Interest in Farmland: Can It Yield Sustainable and Equitable Benefits?," Technical Report 2011.

_ **, Songqing Jin, Yanyan Liu, and Sudhir K Singh**, "Can Labor Market Imperfections Explain Changes in the Inverse Farm Size-Productivity Relationship? Longitudinal Evidence from Rural India," 2016.

Departamendo Nacional de Planeación, "Manual Metodológico para La Determinación de La Unidad Agricola Familiar Promedio Municipal," Technical Report 2000.

Desmet, Klaus and Esteban Rossi-Hansberg, "Spatial Development," *American Economic Review*, April 2014, *104* (4), 1211–43.

Dix-Carneiro, Rafael and Brian K. Kovak, "Trade Liberalization and Regional Dynamics," *American Economic Review*, October 2017, *107* (10), 2908–46.

Eckert, Fabian and Michael Peters, "Spatial Structural Change," Meeting Papers 98, 2018 Society for Economic Dynamics 2018.

Edel, Matthew, "Determinants of Investments by Colombian Community Action Boards," *The Journal of Developing Areas*, 1971, *5* (2), 207–220.

Eswaran, Mukesh and Ashok Kotwal, "Access to Capital and Agrarian Production Organisation," *The Economic Journal*, 1986, *96* (382), 482–498.

Faguet, Jean-Paul, Fabio Sánchez, and Marta-Juanita Villaveces, "The perversion of public land distribution by landed elites: Power, inequality and development in Colombia," *World Development*, 2020, *136*, 105036.

Federación Colombiana de Ganaderos, "Cifras de Referencia del Sector Ganadero Colombiano," Technical Report 2018.

Fernández, Manuel, Ana María Ibáñez, and Ximena Peña, "Adjusting the Labour Supply to Mitigate Violent Shocks: Evidence from Rural Colombia," *Journal of Development Studies*, 2014, *50* (8), 1135–1155.

Foster, Andrew and Mark Rosenzweig, "Economic Development and The Decline of Agricultural Employment," in "Handbook of Development Economics," Vol. 4 2008, pp. Chapter–47.

Foster, Andrew D, "Creating Good Employment Opportunities for the Rural Sector," 2011.

Foster, Andrew D. and Mark R. Rosenzweig, "Are There too Many Farms in the World? Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size," 2017.

Gáfaro, Margarita, Ana María Ibáñez, and David Zarruk, "Equidad y Eficiencia Rural en Colombia: Una Discusión de Políticas Para El Acceso a La Tierra," in Armando Montenegro and Marcela Meléndez, eds., *Equidad y Movilidad Social: Diagnósticos y propuestas para la transformación de la sociedad colombiana*, 2014, chapter 11.

García, Julián, "Evolución de la distribución de las fincas cafeteras Hacía una regionalización de la caficultura colombiana," *Ensayos sobre Economía Cafetera, Federación Nacional de Cafeteros*, 2003, *19*, 193–213.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, August 2020, *110* (8), 2586–2624.

Haggblade, S., P.B.R. Hazell, and T. Reardon, *Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World* International Food Policy Research Institute Series, Johns Hopkins University Press, 2007.

Haggblade, Steven, Peter B. R. Hazell, and Thomas Reardon, "Transforming the rural nonfarm economy: Opportunities and threats in the developing world," Technical Report 2009.

Hamann, Franz, Fernando Arias Rodríguez, Jesús Antonio Bejarano Rojas, Margarita María Gáfaro González, Juan Camilo Méndez Vizcaíno, and Andrea Paola Poveda Olarte, "Productividad total de los factores y eficiencia en el uso de los recursos productivos en Colombia," *Ensayos Sobre Politica Economica*, 2 2019, *2019*, 1–54.

Henderson, Heath and Alan G. Isaac, "Modern value chains and the organization of agrarian production," *American Journal of Agricultural Economics*, 3 2017, *99*, 379–400.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Two perspectives on preferences and structural transformation," *American Economic Review*, 2013, *103*, 2752–2789.

_ , _ , and _ , "Chapter 6 - Growth and Structural Transformation," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, pp. 855–941.

Herrmann, Raoul T., "Large-Scale Agricultural Investments and Smallholder Welfare: A Comparison of Wage Labor and Outgrower Channels in Tanzania," *World Development*, 2017, *90*, 294–310.

Ibáñez, Ana and Juan Muñoz-Mora, *The Persistence of Land Concentration in Colombia: What Happened Between 2000 and 2009?* 01 2010.

Ibáñez, Ana María, *El Desplazamiento Forzoso en Colombia: Un Camino sin Retorno Hacia La Pobreza*, 1 ed., Universidad de los Andes, Colombia, 2008.

IGAC, Instituto Geográfico Agustín Codazzi, "Resolucion 2555 de 1988," 1988.

_ , *Atlas de la Distribucion de la Propiedad Rural en Colombia* 2012.

Inoue, Atsushi and Gary Solon, "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, 2010, *92*, 557–561.

Jaramillo, Carlos Felipe, *Crisis y transformacion de la agricultura colombiana, 1990-2000*, Banco de la Republica, 2002.

Jaramillo, Christian and Ana María Ibáñez, "El censo nacional de población: una comparación de metodologías mediante simulaciones de Monte Carlo," *Coyuntura Social*, 2005, (32), 43–64.

Jiao, Xi, Carsten Smith-Hall, and Ida Theilade, "Rural household incomes and land grabbing in Cambodia," *Land Use Policy*, 2015, *48*, 317–328.

Kalmanovitz, Salomon, "Comentarios sobre el libro "Crisis y transformación de la agricultura colombiana, 1990-2000"," *Fondo de Cultura Económica y Banco de La República*, 2002.

—, *Breve Historia Economica de Colombia* 2015.

Kennedy, Edward H., Scott Lorch, and Dylan S. Small, "Robust Causal Inference with Continuous Instruments Using The Local Instrumental Variable Curve," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2019, *81* (1), 121–143.

Lanjouw, Jean and Peter Lanjouw, "The rural non-farm sector: issues and evidence from developing countries," *Agricultural Economics*, 2001, *26* (1), 1–23.

Le Roux, Guillaume, "Comparabilidad de los censos colombianos de 1993 y 2005: cambios en la recolección de información y dificultades en el análisis de las evoluciones intraurbanas en Bogotá," *Cuadernos de Vivienda y Urbanismo*, 2013.

Liao, Chuan, Suhyun Jung, Daniel G. Brown, and Arun Agrawal, "Spatial patterns of large-scale land transactions and their potential socio-environmental outcomes in Cambodia, Ethiopia, Liberia, and Peru," *Land Degradation & Development*, 2020, *31* (10), 1241–1251.

Ma, Meilin, Jessie Lin, and Richard J. Sexton, "The Transition from Small to Large Farms in Developing Economies: A Welfare Analysis," *American Journal of Agricultural Economics*, 2021.

Machado, Absalón, *La Política de Reforma Agraria y Tierras en Colombia* 2013.

Machado, Absalón and Ruth Suarez, *El mercado de tierras en Colombia : una alternativa viable?*, Santafé de Bogotá, Colombia: Tercer Mundo Editores : CEGA : IICA, 1999.

Mallarino, Consuelo Uribe, "Borrón y cuenta nueva: las estadísticas en Colombia se reinventan a sí mismas," *Universitas Humanistica*, 2007, *63* (63).

Mellor, John W, *Agricultural Development and Economic Transformation Promoting Growth with Poverty Reduction*, Palgrave Macmillan, 2017.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding, " Urbanization and Structural Transformation *," *The Quarterly Journal of Economics*, 03 2012, *127* (2), 535–586.

Morten, Melanie and Jaqueline Oliveira, "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," Technical Report 2018.

Nagy, David, "City location and economic development," Technical Report 2016.

Nupia, Oskar Andrés, "Integración espacial en los mercados laborales: evidencia para las regiones colombianas," *Desarrollo y Sociedad*, 1997, (40).

Ocampo, José Antonio, "Marco Conceptual de la Misión para la Transformación del Campo," Technical Report 2014.

Otero-Cortés, Andrea, "El mercado laboral rural en Colombia, 2010-2019," Documentos de trabajo sobre Economía Regional y Urbana 281, Banco de la Republica de Colombia November 2019.

Pacini, David and Frank Windmeijer, "Robust inference for the Two-Sample 2SLS estimator: Appendix," *Economics Letters*, 2016, pp. 1–8.

Pellegrina, Heitor S, "Trade, Productivity, and the Spatial Organization of Agriculture: Evidence from Brazil," 2020.

Perfetti, Juan José, Diego Escobar, Felipe Castro, Beatriz Cuervo, Maritza Rodríguez, José Ignacio Vargas, Sebastián Martínez Bustos, and Sandra Cortés Acosta, "Costos de Producción de Doce Productos Agropecuarios," Technical Report 2012.

Piniero, Maricel-C., "Globalization and Industrialization of Agriculture: Impacts on Rural Chocontá, Colombia.," *Luna Azul*, may 2016, *43*, 468–498.

Pinzón, José Antonio and Jaume Fonti, "Una aproximación al catastro en Colombia," *Revista UD y La Geomática*, 2007, *1* (1), 25–46.

Putterman, Louis, "A modified collective agriculture in rural growth-with-equity: Reconsidering the private, unimodal solution," *World Development*, February 1983, *11* (2), 77–100.

Ranis, Gustav and Frances Stewart, "Rural nonagricultural activities in development: Theory and application," *Journal of Development Economics*, 1993, *40*, 75–101.

Robineau, Ophélie, Martin Châtelet, Christophe-Toussaint Soulard, Isabelle Michel-Dounias, and Joshua Posner, "Integrating Farming and Páramo Conservation: A Case Study From Colombia," *Mountain Research and Development*, 2010, *30* (3), 212–221.

Rubin, Donald B., "Bayesian Inference for Causal Effects: The Role of Randomization," *Ann. Statist.*, 01 1978, *6* (1), 34–58.

Santaeulàlia-Llopis, Chaoran Chen Diego Restuccia Raül, "The Effects of Land Markets on Resource Allocation and Agricultural Productivity," 2021.

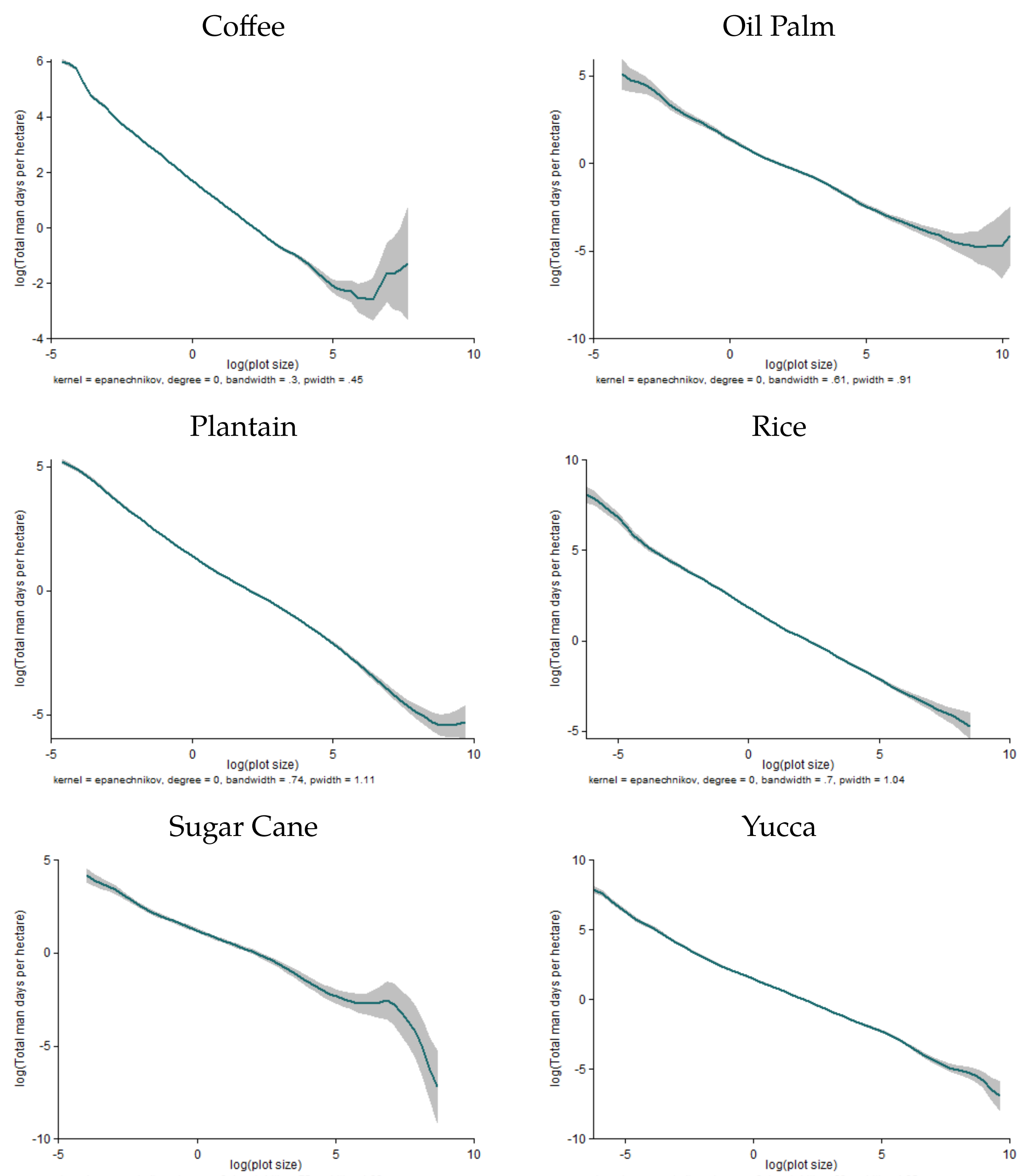
Sen, Abhijit, “Market failure and control of labour power: towards an explanation of ’structure 9 and change in Indian agriculture. Part 1,” *Cambridge Journal of Economics*, 1981, *5*, 201–228.

Sotelo, Sebastian, “Domestic Trade Frictions and Agriculture,” *Journal of Political Economy*, 2020, *128.*

Stock, James H. and Motohiro Yogo, *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press,

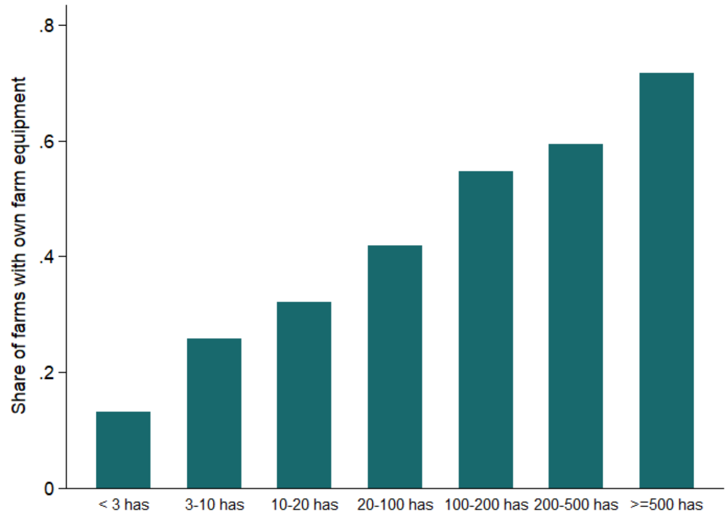
Figures and Tables

Figure 1: Labor Intensity By Plot Size



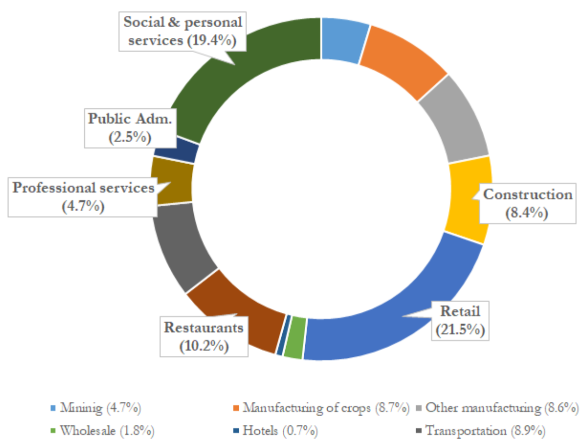
Notes: Own illustrations using data from the 2014 National Agricultural Census.

Figure 2: Equipment Ownership By Farm Size



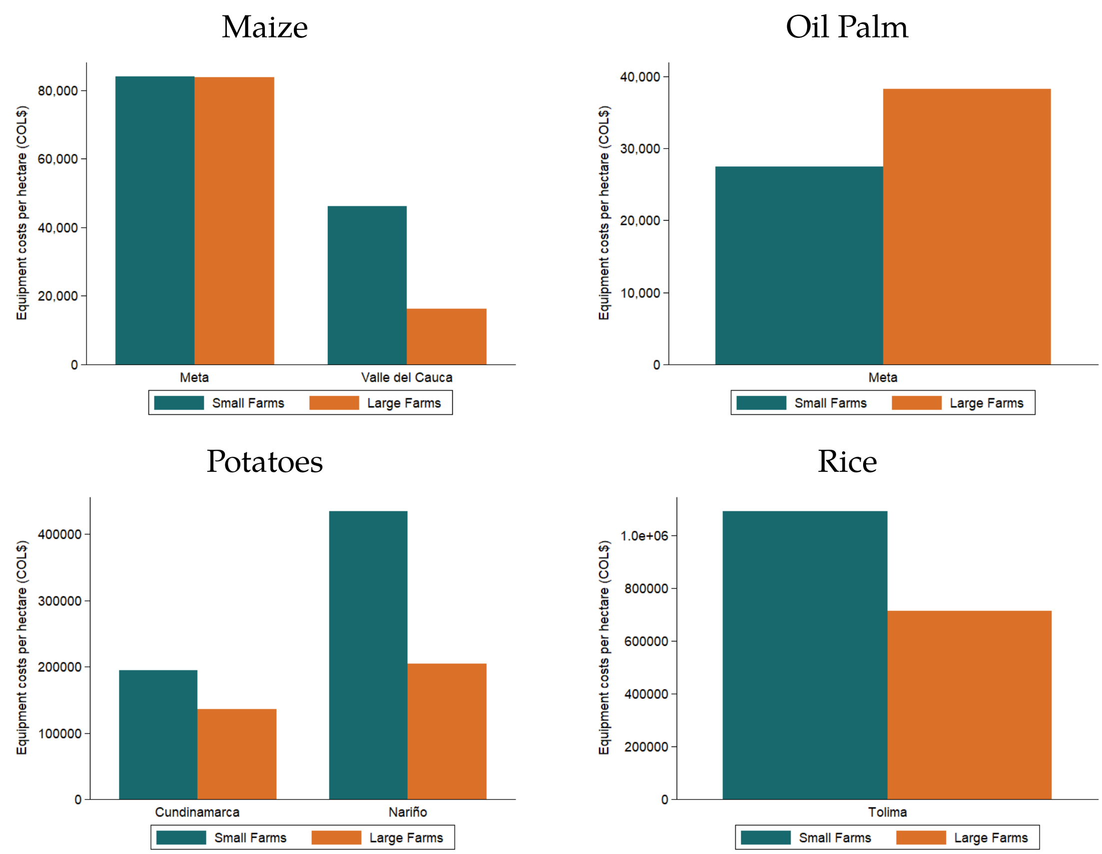
Notes: Own illustrations using data from the 2014 National Agricultural Census.

Figure 3: Distribution of Nonfarm Employment



Notes:. Own illustrations using data from the 2016 National Household Survey (GEIH)

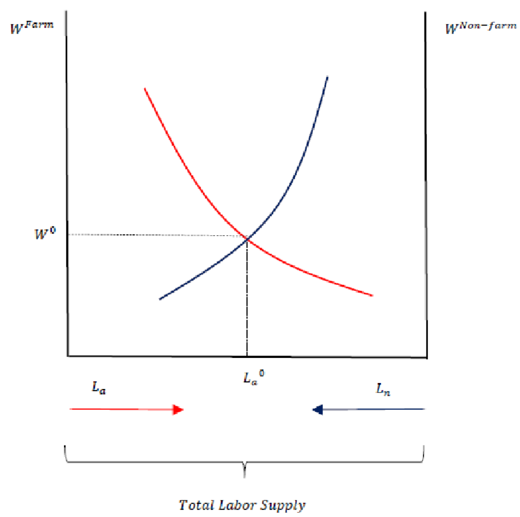
Figure 4: Equipment Costs Per Hectare, By Farm Size



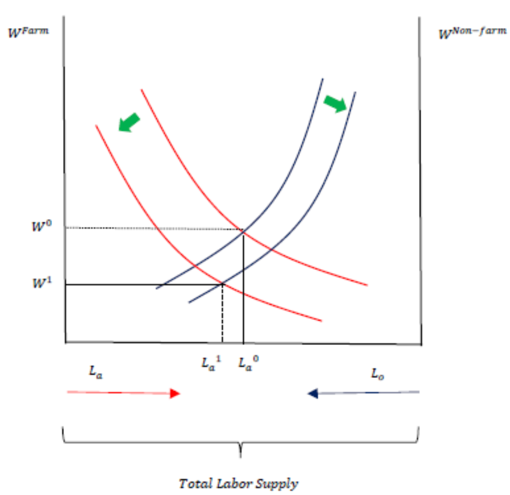
Notes: Own figures based on the data reported in Perfetti et al. (2012). Equipment costs per hectare are measure in Colombian pesos of 2012. Small farms refer to the farms that are at most one family farm unit in size, while large farms refers to farms that are more than one family farm unit in size. For details on the family farm unit measure, see Section 4.1.2.

Figure 5: Overall Effect on Equilibrium Wage: An Illustration

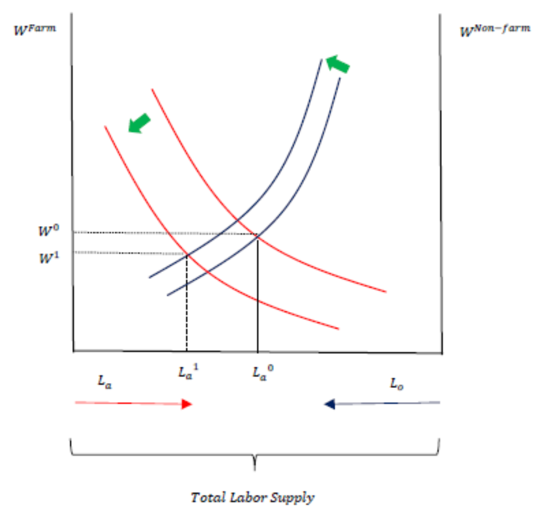
Panel A. The Rural Labor Market



Panel B: Unambiguous Effect

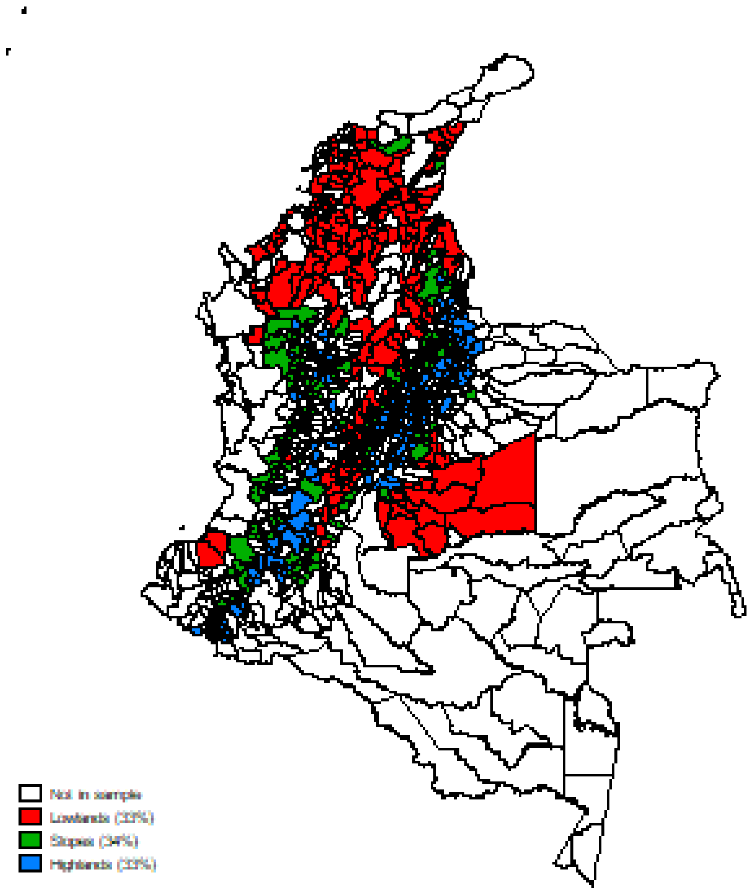


Panel C: Ambiguous Effect



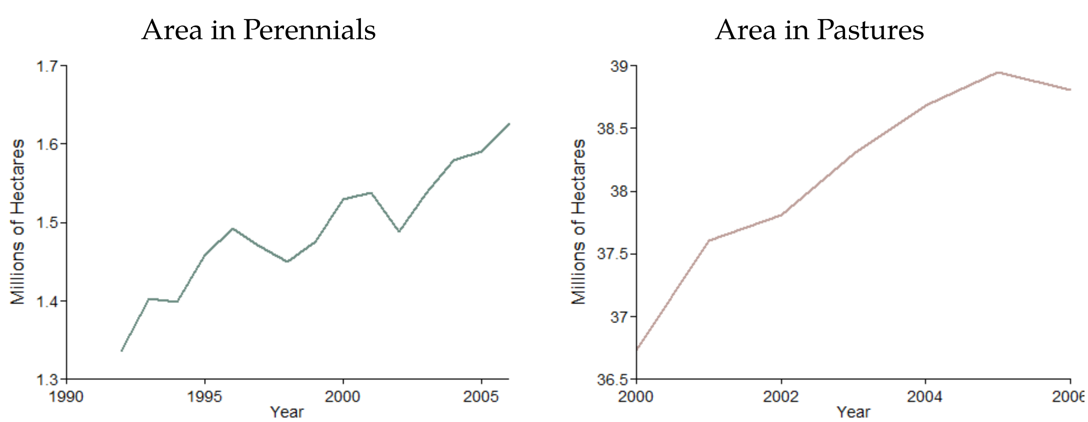
Notes: Own illustration.

Figure 6: Distribution of Municipalities in Study Population by Elevation



Notes: Lowlands are located below 1,000 m.a.s.l., slopes are located between 1,000 and 2,000 m.a.s.l., and highlands are located above 2,000 m.a.s.l.

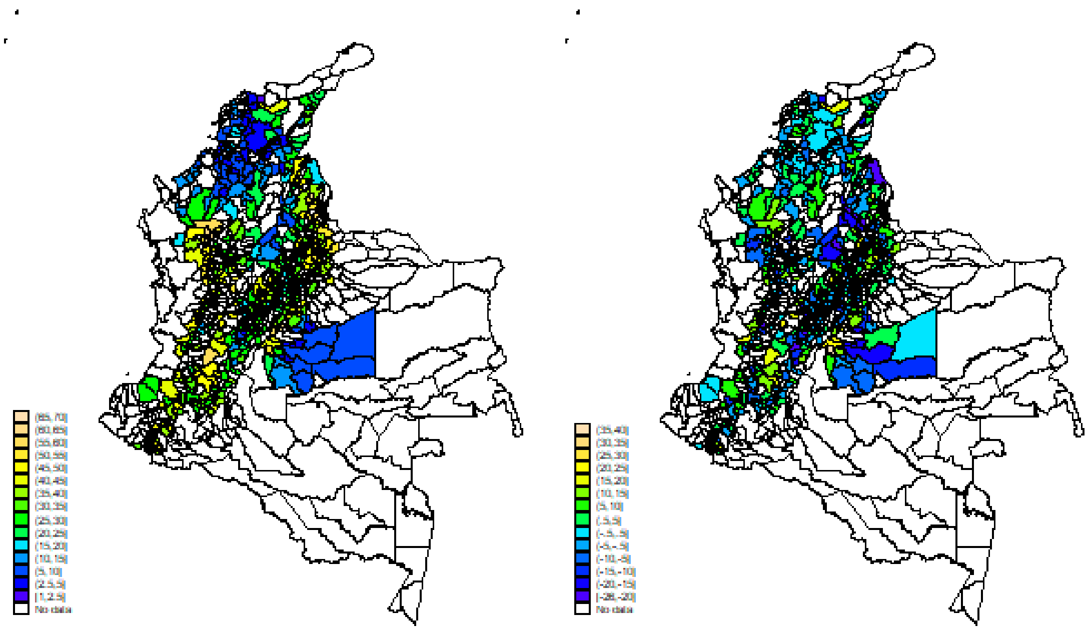
Figure 7: Area in Land-Intensive 1993-2005



Notes: Own illustrations using data from the Municipal Agricultural Evaluations.

Figure 8: Spatial Variation of The Municipality's Slope

Municipality's Slope Municipality's Slope Residuals



Notes: These maps display the spatial distribution of the municipality's slope (in degrees) and its values after partialling out province fixed effects and the change in conflict-related controls 1993-2005.

Table 1: Main Characteristics of The Study Population in 1993

	N	Mean	S.D.	Min	p25	p75	Max
Time-invariant characteristics							
Total area (km2)	590	536.7	1,121.7	20.0	127.0	541.0	17,536.0
Distance to departmen's capital (km)	590	78.7	47.7	9.3	44.2	102.8	276.0
Family farm unit (ha)	590	22.0	14.1	2.5	13.1	27.3	125.6
Time-variant characteristics (1993)							
Total population	590	15,543	14,865	1,277	6,389	19,865	171,936
Share rural population (%)	590	68.7	18.5	4.8	58.4	82.8	97.2
Share of farm employment (%)	590	66.6	17.5	0.7	56.6	80.7	95.9
Unemployment rate (%)	590	2.5	2.2	0.0	1.0	3.5	17.9
Hourly wage (1993 COL$)[†]	1,366	524.1	613.5	3.4	102.2	687.5	6642.9
Gini of landholdings	590	0.38	0.16	0.04	0.26	0.51	0.90
Share of area in large farms (%)	590	28.7	21.4	0.0	11.3	42.4	99.4

Notes: The share of area in large farms is defined, for each municipality, as the share of area in farms that are at least 10 family farm units in size. The statistics on hourly wage refer to the wage workers in the 1998 National Household Survey (†). The average hourly wage of $\$524$ pesos in 1993 corresponds to $\$3,806$ pesos in 2019, which is approximately equivalent to an hourly wage of U$1. See Sections 4.1.1 and 4.2.2 for details on the data sources for the analysis on wages.

Table 2: Incidence of Conflict Between 1993 and 2005

	% Mun.	*Number of events (intensive margin):*			
		Mean	S.D.	p25	p75
Attacks, raids, terrorist acts	95.0%	23.9	32.2	5.0	29.0
Homicides	95.6%	362.6	515.1	54.0	474.0
Share of displaced population†	99.7%	18.9	34.6	1.7	18.4

Notes: The number of municipalities included in this table is 590. Only one municipality did not register any of these conflict-related events during this period. The share of displaced population is calculated with respect to the town's population in 1993 (†).

Table 3: Effect of Large-Scale Concentration on Sectoral Employment & Unemployment Rate (2sls)

	Share of Agricultural Emp.		Unemployment Rate	
	(1)	(2)	(1)	(2)
Share of area in large farms	-0.780***	-0.777***	1.129***	1.118***
	(0.293)	(0.301)	(0.307)	(0.309)
Standardized effect	-0.100	-0.1000	0.145	0.144
Conflict-related covariates		✓		✓
First Stage Results:				
Kleibergen-Paap F-statistic	18.69	17.69	18.69	17.69
Weak-Instrument Robust Inference:				
Anderson-Rubin P-value	0.002	0.004	0.000	0.000
Anderson-Rubin Confidence Set (95%)	[-1.57,-0.36]	[-1.59,-0.35]	[0.69,2.05]	[0.68,2.05]

Notes: The number of municipalities included in these estimations is 590. All the specifications include municipality fixed effects and province-year fixed effects. Standard errors are clustered at the municipality level. The null hypothesis of the Anderson-Rubin test is that the effect of land concentration on the respective outcome is zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Effect of Large-Scale Concentration on Sectoral Employment, & Unemployment Rate (2sls): Different Types of Standard Errors

	Share of Agricultural Emp.			Unemployment Rate		
	(1)	(2)	(3)	(1)	(2)	(3)
Share of area in large farms	-0.777**	-0.777**	-0.777**	1.118***	1.118***	1.118***
	(0.299)	(0.344)	(0.344)	(0.314)	(0.409)	(0.432)
Standardized effect	-0.100	-0.100	-0.100	0.144	0.1441	0.144
Standard errors:						
Cluster municipality level	✓			✓		
Two-way cluster mun & depto-year		✓			✓	
Two-way cluster mun & prov-year			✓			✓
First Stage Results:						
Kleibergen-Paap F-statistic	17.55	11.90	8.75	7.55	11.90	8.75
Weak-Instrument Robust Inference:						
Anderson-Rubin P-value	0.004	0.014	0.009	0.000	0.000	0.000
Anderson-Rubin CS (95%)	[-1.56,-0.33]	[-2.11,-0.26]	[-2.44,-0.26]	[0.67,2.06]	[0.54,2.87]	[0.64,3.65]

Notes: The number of municipalities included in these estimations is 590. All the specifications include municipality fixed effects and conflict-related covariates. The null hypothesis of the Anderson-Rubin test is that the effect of land concentration on the respective outcome is zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Large-Scale Concentration, Sectoral Employment & Unemployment Rate: Pre-Trends Test

	Share Rural Population (Δ_{85-93})		Total Population (Δ_{85-93})	
	(1)	(2)	(1)	(2)
Municipality's slope (degrees)	0.00014	0.00011	-31.39199**	-20.62333
	(0.00028)	(0.00029)	(14.06626)	(14.23055)
Province-Year fixed effects	✓	✓	✓	✓
Conflict-related covariates		✓		✓
Number of municipalities	576	576	576	576

Notes: This table displays the OLS results of the estimation of the change in the share of rural population (and the change in total population) on the municipality's slope before the observed change in land concentration. Standard errors are robust to heteroskedasticity. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Effect of Large-Scale Concentration on Log Hourly Wage (TS2sls)

	(1)	(2)
Share of area in large farms	-2.946	-3.645
	(1.985)	(2.003)
Standardized effect	-0.370	-0.458
Conflict-related covariates		✓
Control variables 1985	✓	✓
First Stage Results:		
Kleibergen-Paap F-statistic	16.857	16.859
Test of equality of coefficients (pval)†	0.670	0.645
Weak-Instrument Robust Inference:		
Anderson-Rubin P-value	0.211	0.120
Anderson-Rubin Confidence Set (95%)	[-8.53,1.69]	[-9.12,0.88]

Notes: The number of observations included in these estimations is 2,529. The first stage is estimated using individuals from 717 municipalities, while the second stage uses individuals from 49 municipalities. All the specifications include municipality fixed effects, department-year fixed effects, and interactions of both the population size and the share of rural population in 1985 with year fixed effects. Standard errors are robust to heteroskedasticity. The second row in *First Stage Results* tests the null hypothesis that the first-stage coefficient is equal across both samples, which is an additional assumption for the consistency of the two-sample 2sls estimator (†). The null hypothesis of the Anderson-Rubin test is that the effect of land concentration on the hourly wage is zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

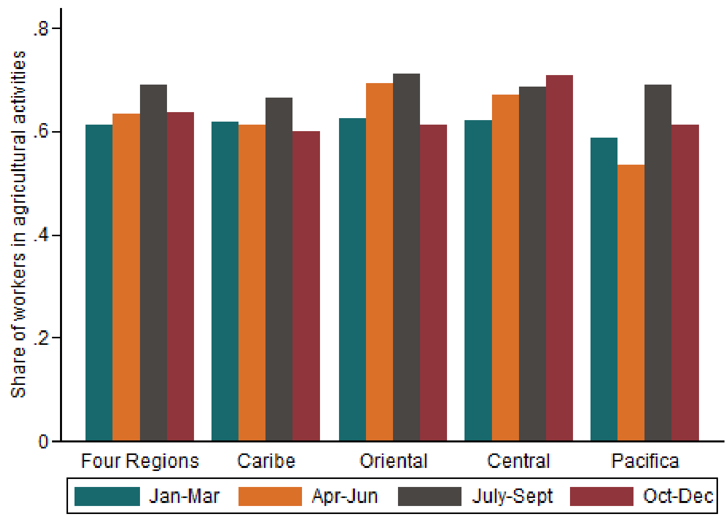
Table 7: Large-Scale Concentration and Log Hourly Wage: Pre-Trends Test

	Share Rural Population (Δ_{85-93})		Total Population (Δ_{85-93})	
	(1)	(2)	(1)	(2)
Municipality's slope (degrees)	0.00010	0.00001	-47.31535***	-35.34544***
	(0.00021)	(0.00021)	(10.80324)	(10.03544)
Department-Year fixed effects	✓	✓	✓	✓
Conflict-related covariates		✓		✓
Number of municipalities	749	749	749	749

Notes: This table displays the OLS results of the estimation of the change in the share of rural population (and the change in total population) on the municipality's slope before the observed change in land concentration. Standard errors are robust to heteroskedasticity. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Appendix Figures and Tables

Figure A1: Share of Agricultural Employment in Rural Areas, by Quarter and Region



Notes: Own illustrations using data from the National Household Surveys.

Table A1: Correlation Between Instrumental Variable & Conflict-Related Covariates

	Log (Attacks, raids, etc)		Log Homicides		% Disp. Pop. (Exp.)		% Disp. Pop. (Rec.)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Instrumental Variable	0.0046***	0.0003	0.0033***	-0.0004	0.0004***	0.0004***	0.0002***	0.0002**
	(40.05)	(1.11)	(34.10)	(-1.57)	(11.82)	(3.48)	(9.98)	(2.84)
Municipality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Province-Year fixed effects		✓		✓		✓		✓
Number of municipalities	590	590	590	590	590	590	590	590

Notes: This table displays the OLS correlations between the instrumental variable (i.e. interaction of the municipality's slope and year) and four conflict-related covariates: i) the sum of attacks, raids, and terrorist acts by all perpetrators (in logs), ii) the logarithm of the number of homicides, iii) the share of displaced population that left the town, and iv) the share of displaced population that was received in the town. Standard errors are clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: Large-Scale Concentration, Sectoral Employment, & Unemployment Rate: The Importance of Province-Year Fixed Effects (2sls)

	Share of Agricultural Emp.			Unemployment Rate		
	(1)	(2)	(3)	(1)	(2)	(3)
Share of area in large farms	1.198***	-0.574**	-0.752**	-0.650***	1.017***	1.118***
	(0.169)	(0.278)	(0.299)	(0.100)	(0.314)	(0.314)
Standardized effect	0.154	-0.074	-0.0967	-0.084	0.131	0.144
Department-Year fixed effects		✓			✓	
Province-Year fixed effects			✓			✓
First Stage Results:						
Kleibergen-Paap F-statistic	87.61	17.46	17.55	87.61	17.46	17.55
Weak-Instrument Robust Inference:						
Anderson-Rubin P-value	0.000	0.029	0.004	0.000	0.000	0.000

Notes: The number of municipalities included in these estimations is 590. All the specifications include municipality fixed effects and conflict-related covariates. Standard errors are clustered at the municipality level. The null hypothesis of the Anderson-Rubin test is that the effect of land concentration on the respective outcome is zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Preliminaries of Quantitative Analysis

B.1 Calibration

I calibrate the model based on the Colombian rural economy. The parameters to calibrate are technological parameters, preference parameters, distributional parameters, and endowments. The calibration strategy follows two steps. First, I calibrate some parameters externally based on values taken from the literature. Some of them are assigned directly, and others are set independently of the model's equilibrium outcomes. Then, I calibrate the remaining parameters using the model's structure to match targeted moments at the aggregate and farm level.

Parameters calibrated externally: Table B3 summarizes the parameters that are calibrated based on values from the literature. I normalize productivity and the output price in agriculture (A, p_a) to 1 and set $\gamma = 0.40$ based on the land shares estimated by Avila and Evenson (2010).[58] The relative weights of consumption goods in preferences are set based on the empirical estimates of the structural transformation literature (Herrendorf et al., 2013). In particular, I set ω_a equal to the weight of agricultural consumption and ω_n to the weight of consumption in services. The motivation for using estimates from the service sector to calibrate ω_n is that the local nonagricultural good in the model is non-tradable.

Based on the estimates of Gáfaro et al. (2014), I set the share of landed agents in the economy equal to 0.42. In addition, I assume that the distribution of land among these

[58]This parameter value is obtained as the area-weighted average of the land share in crop production (0.23) and livestock raising (0.44) in Colombia. The land share for crops is close to that of Peru (Sotelo, 2020), a country with a similar crop portfolio, and is smaller than Brazil's share (Pellegrina, 2020), characterized for production at a larger scale.

agents is Pareto, which is a standard way to model land heterogeneity in the development literature: $F(h) = 1 - \left(\frac{h_m}{h}\right)^{\tau}$.[59] Given this distribution, I set $\tau = 1.07$ and $h_m = 1.09$ to match a Gini coefficient of 0.88 and an average farm size of 16.9 hectares.[60] Using the agricultural census, I estimate the total number of farmland hectares at $H = 34.2$ million. Thus, to match a farmland-to-labor ratio of 7.03, I set $L = 4.863$ million.[61] This latter value is in line with recent rural workforce estimates using the national household surveys (Otero-Cortés, 2019). Finally, I set $r = 7\%$, which is the average real interest rate in the Colombian economy over the last decade.

Table B3: Parameters Calibrated Externally

Description	Paramater	Value
Productivity in agriculture	A	1.000
Price agricultural good	p_a	1.000
Land share of income	γ	0.398
Weight agricultural consumption	ω_a	0.020
Weight local nonagricultural consumption	ω_n	0.810
Share of landed agents	θ	0.416
Pareto distribution (shape)	τ	1.069
Pareto distribution (scale)	h_m	1.091
Aggregate farmland	H	34,207
Aggregate labor	L	4,863

Notes: The units for aggregate farmland and labor are thousand of hectares and individuals, respectively.

Parameters calibrated using the structure of the model: Five parameters are jointly calibrated to match five aggregate moments from the Colombian data. The calibrated parameters are the relative importance of labor to capital, the determinants of the elasticity of substitution and capital pricing, and the non-homothetic parameters. Table B4 summarizes the parameter values and the moments used in the calibration. For each parameter, I also present the moment's target value and its respective model counterpart. I discuss each of these in turn.

The moments associated with the non-homothetic parameters $\bar{c_a}$ and $p_c\bar{c_c}$ are measured using the rural component of the National Households' Budget Survey.[62] The targeted moments are the aggregate share of final expenditure in agricultural goods and nonagri-

[59] See, for instance, Bazzi (2017), Carter and Kalfayan (1989), and Eswaran and Kotwal (1986).

[60] This farm size value is calculated using data from the agricultural census and is similar to the one estimated by Hamann et al. (2019).

[61] Given the Pareto distribution, the farmland-to-labor ratio is $\frac{H}{L} = \theta\frac{\tau h_m}{(\tau-1)}$.

[62] The quantitative exercise discussed in this section does not require the identification of $\bar{c_c}$ and p_c separately. Further, since prices are the same for everyone, expenditure in consumption is proportional to quantity consumed. To identify $p_c\bar{c_c}$, I use the following transformation of the preferences: $u(c) = \omega_a log(c_a - \bar{c_a}) + \omega_n log(c_n) + \omega_c log(p_c c_c + p_c\bar{c_c}) - log(p_c)$.

cultural tradable industries.[63]

In addition to these moments, I select the parameter that determines the nonlinearities in capital pricing ν to match the ratio of capital costs between large and small farms. This ratio is calculated using data from the study conducted by Perfetti et al. (2012), which reports production costs by farm size for four main crops in the country: oil palm, potatoes, rice, and maize.[64] The target value of 0.86 is calculated as the area-weighted average across all crop-specific ratios. Finally, the relative importance of labor to capital in production η is selected to match the labor cost share in agriculture estimated by Avila and Evenson (2010). This target is calculated as the area-weighted average of the labor cost share for crops and livestock farming and takes a value of 0.36.

Table B4: Parameters Calibrated Using The Structure of The Model

Paramater	Value	Moment	Target	Model
η	0.871	Labor cost share	0.36	0.14
ρ	0.985	Ratio of labor per hectare (smallest/largest farm)	240	239.6
ν	0.171	Capital cost ratio (large/small farms)	0.77	0.99
$\bar{c}_a$	0.900	Expenditure share in agriculture	0.31	0.35
$p_c\bar{c}_c$	1.64E-05	Expenditure share in nonagricultural tradables	0.13	0.12

Notes: The target value for the labor cost share is taken from Avila and Evenson (2010). The expenditure shares are calculated using the households' budget survey, labor per hectare is calculated using the agricultural census, and the capital cost ratio is calculated using data from Perfetti et al. (2012).

[63] The tradable aggregate expenditure share is computed using the total monthly household expenditure in restaurants and hotels, home furniture, alcoholic beverages, clothing, and footwear.

[64] These four crops account for 42% of the national cultivated area. A future version of this draft will include cost data from coffee production, which is the crop with largest cultivated area at 16%.