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

Karen Ortiz-Becerra[†]
UC Davis

November 15, 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 land 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 non-homothetic 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. I examine this question in the Colombian setting by assembling a novel dataset and leveraging quasi-experimental variation in the ability of rural counties to respond to a trade shock that changed land use. Regions with an increase in large-farm consolidation experienced a decline in the share of agricultural labor and a sizeable increase in unemployment rates. These findings shed light on the distributional impacts of consolidation across individuals and the potential implications of structural transformation within rural economies.

[†]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 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 agricultural land 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 two 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 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 redistribute

[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). Specialty and certified coffee, for instance, are often more profitable when produced at a small scale.

[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.

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 common to agricultural production models – differences in factor intensity across the farm size distribution and efficiency gains from scale – with a third feature frequently observed in rural economies – the non-homothetic demand growth for local goods. A key feature of the model is that land consolidation simultaneously affects the demand for labor in both economic sectors.

Using this framework, I show that consolidation leads to a push response in the demand for labor out of the farm sector and a likely pull response in nonfarm labor demand. On net, if the pull response is small relative to the push response in the farm sector, consolidation may lead to a reduction in wages, which is the opposite relationship than we usually associate with the process of structural transformation. These theoretical results highlight the importance of local multiplier effects to generate employment gains after consolidation. This is especially relevant in developing contexts where labor mobility barriers limit workers' reallocation across space. (Morten and Oliveira, 2018; Dix-Carneiro and Kovak, 2017; Bryan et al., 2014).

Next, I empirically 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 using measures of wages and unemployment rates.

To estimate treatment effects, I leverage quasi-experimental variation in the ability of local economies to respond to a common 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, the upsurge in conflict in rural areas, which had potential differential effects across economies in the lowlands and the slopes (Centro Nacional de Memoria Historica, 2013). Second, trade-induced labor market trends that are associated with the production of specific crops. For instance, changes in tariffs and relative prices are likely to benefit economies that produce exportable crops and put at disadvantage those that produce importable ones (Dix-Carneiro and Kovak, 2019, 2017). In this analysis, the empirical strategy then consists in comparing municipalities with similar agricultural suitability but different abilities to consolidate. Thus, identification comes from the variation in the physical ability of production at scale rather than the suitability of a local economy to produce a particular crop.

I find that the consolidation of land in large farms prompted the reallocation of labor out of agriculture within local economies. Specifically, a one standard deviation increase in consolidation resulted in a ten percentage point decrease in the share of farm workers. In addition, consolidation led to a 14 percentage point increase in the unemployment rate with no significant change in wages. This sizeable increase in unemployment suggests that the net impact on workers' income was negative. My estimates imply that one standard deviation increase in land consolidation explains 70% of the observed decline in farm employment for the *average* Colombian municipality between 1993 and 2005. This increase in consolidation is equivalent to a twelve percentage point rise in the share of land in large 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 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.

Taken together, my findings indicate that consolidation led to a decrease in the demand for farmworkers 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. While I cannot directly test for mechanisms, descriptive results corroborate the importance of non-homothetic consumption growth, as the absolute decrease in labor demand appears to be smaller in economies where local multiplier effects are expected to be large. My results on unemployment suggest that labor mobility was not large enough to offset the decrease in labor demand after consolidation.[4] These findings are consistent with the low levels of integration across rural labor markets during the 1990s (Nupia, 1997) and support previous empirical results on the delayed adjustment of labor markets to economic shocks (Dix-Carneiro and Kovak, 2019; Dix-Carneiro, 2014; Autor et al., 2014; Artuç et al., 2010). For instance, Dix-Carneiro and Kovak (2019) find that trade liberalization's effects on unemployment and informality in Brazil persisted over ten years.

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.[5] 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 wages and the labor reallocation across sectors, and assess how much of these effects are driven by non-homothetic local consumption growth. To evaluate impacts in the short and long term, I will quantify these effects under different levels of labor mobility across space.

This paper contributes to the long-standing 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; Foster and Rosenzweig, 2017; Henderson and Isaac, 2017; Eswaran and Kotwal, 1986; Carter and Kalfayan, 1989), and macroeconomic models that quantify the contribution of misallocation to productivity gaps (Santaeulàlia-Llopis, 2021; Adamopoulos and Restuccia, 2020; Adamopoulos et al., 2019; Chen, 2017; Adamopoulos and Restuccia, 2014). My work complements these studies by examining implications for local labor markets and adds to this literature by introducing a framework that considers the links between the scale of production and labor demand in both the farm and nonfarm sectors. This framework combines insights from the relationship between farm size and productivity

[4]This is not to say that the extent of rural-out migration was small during this period. On the contrary, due to the upsurge of conflict, a large proportion of rural households were forced out of rural communities. In my analysis, however, forced displacement is a potential threat to internal validity. Thus, my empirical strategy compares municipalities with a similar incidence of this type of labor mobility. Future access to net migration flows will allow me to assess how much of the observed changes are driven by economic migration.

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

with insights from the literature on the rural nonfarm economy that had not yet been formalized. Namely, the importance of local consumption on employment in the nonfarm sector (Foster, 2011; Haggblade et al., 2009, 2007; Foster and Rosenzweig, 2008; Lanjouw and Lanjouw, 2001), and the connection between income concentration and labor demand (Mellor, 2017; Ranis and Stewart, 1993). By embedding these insights into a general equilibrium framework, I show that land consolidation might have different impacts across individuals and could negatively affect rural workers despite the productivity gains.

This paper also contributes to the growing empirical work that examines the effects of large-scale transactions on the welfare of rural populations (Liao et al., 2020; Ali et al., 2019; Deininger and Xia, 2016). With a few exceptions, this body of work consists of descriptive papers or studies that rely on cross-sectional data and unconfoundedness to estimate spillover effects on smallholders (Bottazzi et al., 2018; Herrmann, 2017; Jiao et al., 2015). I contribute to this literature by providing novel empirical evidence on the economy-wide impacts of large-scale consolidation on aggregate employment and wages.

A second line of work related to this paper is the literature on structural transformation (Herrendorf et al., 2014). My empirical approach relates to previous work that exploits quasi-experimental variation to examine the determinants of structural change at a local level rather than across regions (Bustos et al., 2020; Uribe Castro, 2020; Bustos et al., 2016; Foster and Rosenzweig, 2008). My empirical work relates to the research on the impacts of trade on structural transformation (Farrokhi and Pellegrina, 2020; Fajgelbaum and Redding, 2018; McArthur and McCord, 2017), including Laskievic (2021) who finds that commodity price booms affect labor reallocation through greater land inequality and agricultural input use. However, while these papers study the direct effect of trade leveraging the differential exposure of municipalities to shocks, I study the direct impact of land consolidation by exploiting the isolated effect of trade on land concentration across towns with similar exposure to shocks. Using this approach, I find that land consolidation is another driver of the reallocation of labor out of agriculture in local economies.

Finally, my findings also indicate that the reallocation of labor out of the farm sector can occur along with a decrease in workers' income. This implies the opposite relationship than we usually observe with the process of structural transformation. These results are related to recent studies that examine the spatial implications of structural transformation (Bustos et al., 2020; Eckert and Peters, 2018; Nagy, 2016; Desmet and Rossi-Hansberg, 2014; Michaels et al., 2012; Caselli and Coleman II, 2001). In particular, Eckert and Peters (2018) also quantify a reduction in rural income per capita after the transformation of the US economy during the 20th century. In line with their work, my empirical results shed light on the prevalence of barriers to geographic labor mobility and are consistent with findings in the literature that investigate the capacity for labor markets to respond to economy-wide shocks beyond the short-term (Dix-Carneiro and Kovak, 2019; Dix-Carneiro, 2014; Autor et al., 2014; Artuç et al., 2010).

The rest of this paper is organized as follows. In the next section, I establish empirical patterns that relate large-farm consolidation and employment in each sector. Section 3 develops the model and discuss its main analytical insights. In Section 4, I describe the empirical setting and examine the impact of large-farm consolidation during the 1990s in Colombia. Section 5 concludes.[6]

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.[7] By contrast, Figure 2 shows that large 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

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

[7]This relation remains practically unchanged when using the total area cultivated as opposed to the total farm size. It also holds when using the total number of permanent workers as a measure of employment instead of the total man days.

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.[8] 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.[9] 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.[10] These advantages in equipment markets render production at scale more profitable. Thus, to the extent that these profits are used for local consumption, 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

[8] 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.

[9] 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 $\frac{p_s}{p_l} > \frac{k_l h_s}{k_s h_l} > 1$.

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

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_n, 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. Consistent with the patterns documented in Section 2, there are pecuniary economies in the agricultural capital market. Namely, 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.[11]

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 $\rho \rightarrow 1$, labor and capital become highly substitutable.

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

$$Q_n = L_n^{\alpha} K_n^{1-\alpha}$$

The nonfarm good is nontradable and thereby production depends on local demand. This characterization precludes the production of value-added products to focus attention on the role of local consumption as a driver of employment in this sector (see Section 2).[12] In contrast to agricultural production, there are no pecuniary economies in the production of the nonfarm good. Thus, the effective price of capital for the stand-in firm equals the equilibrium rate r and $\pi_n^* = 0$.

[11]See Section 3.4 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).

[12]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 represents a two percentage point increase since 2006.

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 $\bar{c_c} > 0$.[13] 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.[14]

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, $Y = w + \pi_a^*$.

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 $0 < \nu < 1$.[15]

[13] 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.

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

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

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 non-farm 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 (do not) purchase the city good.[16] Since $\pi_n^* = 0$, p_n^* equals the marginal cost $MC(w,r)$ in equilibrium.

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 sector L_n and the total workers required by the farmers L_a. Since the supply of the non-farm good is perfectly elastic, L_n is pinned down by local consumption C_n.

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

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

where K is the total supply of capital equipment in the economy, K_n is the amount of capital used in the nonfarm sector, and K_a is the aggregate demand for capital in the farms.

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

Definition 1. *The equilibrium in this economy is given by input demands* $\{L_a, L_n, K_a, K_n\}$*, 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 $p_n^* = MC(w, r)$.
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 price p_a^* is set internationally and p_c^* is set in city.

3.3 Model Characterization

3.3.1 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.

3.3.2 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, $\mathcal{C}_n$, 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.

3.4 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.[17] Thus, we can obtain a measure of land valuation for each agent i as

[17]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.

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 Zimmerman, 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

This section examines the impact of a specific shift in land concentration during the 1990s in Colombia. This shift was largely driven by the market integration with the World's economy, which increased the area in land-intensive crops. I first introduce the data and context. Then, I describe the empirical strategy. Finally, I discuss the results and their implications in light of the model in Section 3.

4.1 Data Sources and Study Population

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.[18] I focus on rural economies and exclude large cities and their main agglomerations using the definition proposed by the *Rural Mission* in 2014.[19] To account for the segregation and creation of new municipalities across time, 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.

Unemployment and sectoral employment: I use information from the National Population Census of 1993 and 2005 to construct these measures. 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 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 construct the share of agricultural workers using information on the economic sector of the main job in the last 15 days. 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.[21]

Wages: Since the population census does not collect data on wages, I use individual-level data from the National Household Surveys (ENH) of 1998 and 2009. These repeated cross-sections 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 mod-

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

[19] This definition classifies municipalities in four categories based on population density and the number of inhabitants living in the town's seat. See Ocampo (2014) for more details.

[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). Access to this data was obtained per confidential agreement with the National Department of Statistics.

[21] Two main reasons indicate that the timing of data collection does not primarily drive the differences in agricultural employment across locations. First, in most regions, the share of agricultural workers does not vary much throughout the year (see Figure A1 in the appendix). Second, according to experts at the National Department of Statistics, the roll-out of the censuses was not correlated with any particular season.

ule of this survey collects data for a random sample of workers on features such as occupation, hours worked, and income. 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. For all outcomes, I focus on individuals between 15 and 65 years old.

Land consolidation: I use data from the national cadastre system for the years 1993 and 2005 to construct measures of land consolidation.[22] 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] For each municipality and year, I have information on the number of properties and their total area across thirteen size ranges.[24] To focus on privately held land, I exclude the records of indigenous reserves and properties that belong to the State. Two main features make this data advantageous to construct measures of concentration and consolidation. First, the system records information based on possession instead of ownership. Therefore, the records of private land include farmers with informal titles or settled in vacant public lots. Second, the measures of plot size are less likely to be afflicted by self-reporting bias – common in survey data– as the data gathered upon the *cadastre formation* is collected and updated by personnel in the field (IGAC, 1988).[25]

I define large-farm consolidation as the share of area in large farms. To determine what a large plot is for each town, I use as reference the average family farm unit, a policy instrument representing the minimum plot size needed to generate an income surplus given the agro-ecological conditions of the plot's location.[26] Following Machado and Suarez (1999), I use a threshold of ten units to characterize large plots. Therefore, large-farm consolidation for each municipality i is given by: $\omega_i^{large} = \omega_i^a \quad if \quad \kappa_i \in [a, b)$, where κ_i refers to ten

[22] The data from 2005 was purchased from the National Institute of Geographic Information (IGAC), and the data from 1993 was generously provided by Fabio Sánchez at the University of the Andes in Colombia.

[23] The national cadastre is managed by five different agencies: the National Institute of Geographic Information (IGAC), the department of Antioquia, and the capital cities of Bogota, Medellin, and Cali (IGAC, 2012). In this analysis, I use the information from Antioquia and IGAC, which constitutes the whole universe of rural municipalities with cadastral data.

[24] These size ranges 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.

[25] One important caveat of this data, however, is that resource limitations prevent 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, to the extent that the updates are not correlated with municipality's characteristics in general (Martinez, 2020), it is unlikely that my estimates are largely driven by differences in measurement error.

[26] The family farm unit was initially created by the Law 135 of 1961 to guide the allocation of vacant public lands. It represents the minimum plot size required to produce an income of three monthly minimum wages and a disposable income after paying land rent payments. This unit takes on different values depending on the type of agro-ecological zone and land use. Thus, the average unit for each town is calculated as the weighted mean across production systems and zones (Departamendo Nacional de Planeación, 2000).

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 a hectares.[27] In addition to this measure, I also use this data to calculate the Gini index of landholdings, which is a measure of *overall* concentration across the whole plot distribution.

Additional municipal features: I compile a set of municipality characteristics using different data sources. First, I obtain information on administrative divisions and the town's average family farm unit from the National Department of Statistics (DANE). Second, I use administrative data from the National Memory Center to calculate measures of conflict intensity and the prevalence of forced displacement between 1993 and 2005.[28] Third, I digitize data on initial levels of urbanization, such as the total number of inhabitants and the share of the rural population, using the 1985 census. Finally, I calculate topographic measures on elevation and the average degree of terrain's inclination using the Data Elevation Model from NASA and the shapefiles of municipalities with 1993 boundaries. In contrast to other sources, these topographic measures are very precise as the pixel resolution is one arc-second (approximately 30 meters). This is particularly relevant for my analysis since it implies that the instrument I construct for land consolidation displays a considerable variation across towns.

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.[29] My study sample accounts for three-fourths of the rural economies in these departments.[30] These towns are spread in similar proportions across lowlands, hills, and highlands, thereby representing the main agro-climatic regions in the country.

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 are located about 79 kilometers from the department's capital and have a mean area of 537 square kilometers. In contrast to large cities, the population count was low and

[27]For instance, if the average family farm unit for the town is 32 hectares, my measure of large-farm consolidation 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, I plan to check whether the main results of the paper are robust to different definitions of ω_i^{large} in a future version of this draft.

[28]These indicators include the share of the forcefully displaced population, the number of murdered individuals, and 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).

[29]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: i) department was not part of the agricultural frontier in the 1990s, ii) rural areas consisted mostly of indigenous reservations and afro-colombian lands, iii) cadastral data was either non-existent or unreliable.

[30]The remaining 25% of rural towns do not have complete data for both years of analysis.

the the majority of workers worked in agricultural activities. These rural towns also had a low unemployment rate and a wide spread in the wage perceived by workers. While the average hourly wage was slightly above the minimum legal wage at COL$524, close to two-thirds of the workers were making less than that amount.[31]

Finally, the size threshold that defines a large farm has a mean of 220 hectares and varies widely across towns. For instance, while in some municipalities, a large farm has at least 800 hectares, in others, farms of 25 hectares are considered large. The share of area in large farms across municipalities was 29% on average and one-quarter of these towns had a share above 42%.

4.2 Trade Liberalization and Shift in Land Concentration

Colombia underwent a process of trade liberalization in the early 1990s. This liberalization ended an import substitution regime that lasted over forty years and was characterized by the gradual reduction of import tariffs, the enactment of several trade agreements, and the implementation of policies to promote exports. The average tariff in the whole economy went from 38.6% to 11%, and from 31.5% to 15% in the agricultural sector (Jaramillo, 1998). After the integration, the agricultural sector experienced a productive transformation and a change in land use. The area in pastures and cash crops increased by more than two million hectares, while the area to grow importable crops decreased (see Figure 6).

This productive transformation was accompanied by a general shift in land concentration. In some regions, farms were consolidated to produce land-intensive cash crops such as cattle, flowers, and oil palm.[32] In others, farms were fragmented to mitigate the impacts of the deterioration in the terms of trade. Overall, rural municipalities experienced an increase in the Gini index of landholdings during this period, and large-farm consolidation changed over 90% of the towns.[33] Large-farm consolidation displayed a marginal decrease over this period and there was a wide variation in the direction and magnitude of this change across locations (see Figures 7 and 8).

4.3 Identification Strategy

To examine the impacts of large-farm consolidation, I use a 2sls approach that exploits the differential ability of the municipalities to respond to this trade-induced shock. In par-

[31] An hourly wage of COL$524 in 1993 is equivalent to COL$3,806 per hour in 2019 (U$1 per hour).

[32] These crops are usually produce at large-scale due to a high degree of vertical integration. Furthermore, cattle has a maximum carrying capacity per hectare that requires a minimum plot size threshold to render production profitable.

[33] Laskievic (2021) also finds an increase in the Gini index after the commodity price boom in Brazil during the 2000s.

ticular, I construct an instrument for the change in consolidation based on topographic features of the economy that influence the financial viability of production at a large scale. This instrument is defined as the average degree of inclination for the terrain in each municipality.[34] In my sample, the degree of inclination ranges from $1.5°$ to $64.8°$ and has an average of $29.7°$.

The rationale behind this instrument is that the construction and maintenance of infrastructure for large-scale production are cheaper in flatter terrains. For instance, export crops such as flowers and oil palm require greenhouses and irrigation systems. Likewise, flat landscapes are more attractive for mechanization and extensive cattle ranching, as the carrying capacity per hectare decreases with the gradient of inclination.[35] Declines in large-farm consolidation are also smaller in locations with lower degree of inclination, as the opportunity cost of fragmentation is higher. For instance, while coffee farms in the slopes are fragmented after declines in the international coffee price, coffee farms in the inter-Andean valleys are often transformed into cattle farms (Balcázar, 2003; García, 2003). This negative relationship between large-farm consolidation and the economy's degree of inclination is depicted in Figure 9.

A town's topography also influences crop suitability. Thus, one important threat to the exclusion restriction of this instrument is that locations with different crop portfolios faced distinct income and labor market shocks due to trade liberalization. For instance, economies that were suitable to produce cash crops experienced an improvement in their terms of trade, while economies that produced importable crops were negatively affected in terms of profitability. To address this concern, I compare contiguous municipalities of the same province (δ_p), which share similar crop suitability and are subject to the same governmental policies.[36] Similarly, since mountains provide a natural shelter for illegal groups (Centro Nacional de Memoria Historica, 2013), the upsurge in conflict during this period likely had differential effects on labor markets of economies with distinct topography. I account for this potential relationship by including measures on trends in forced displacement and conflict-related events as covariates (ΔC_{mp}).

For the analysis on unemployment and sectoral employment, I estimate equations of the form:

[34]This measure corresponds to the area-weighted average across all the raster cells that intersect with the municipality's surface. For more details on the data used to construct this measure, see Section 4.1.

[35]According to interviews in the field, the carrying capacity of cattle in flatlands is about three cows per hectare compared to two cows (or less) in the slopes.

[36]Provinces are department subdivisions that have been historically used to plan and develop environmental and territorial policies (DANE, 2014). There are 100 provinces in my sample with an average of eight municipalities per province.

$$\Delta Y_{mp} = \beta \Delta \hat{D}_{mp} + \delta_p + \alpha_1 \Delta C_{mp} + \Delta u_{mp}$$
$$\Delta D_{mp} = \gamma S_{mp} + \delta_p + \alpha_2 \Delta C_{mp} + \Delta \eta_{mp} \tag{6}$$

where ΔY_{mp} refers to the change in the outcome of interest between 1993 and 2005 for municipality m in province p, ΔD_{mp} refers to the change in large-farm consolidation, and S_{mp} refers to the terrain's degree of inclination. In all estimations, standard errors are clustered at the municipality level to account for serial correlation. To give an idea of the variation I use to estimate β, I show the spatial distribution of the instrument and its values after partialling out the fixed effects and covariates in Figure 10.

For the analysis on wages, I use repeated cross-sections of workers in rural municipalities. Thus, I estimate the following modified version of Equation 6:

$$Y_{imjt} = a_m + \beta \hat{D}_{mjt} + \delta_{j,t} + \alpha_1 C_{mjt} + \lambda_1 X^{85}_{m,t} + u_{imjt}$$
$$D_{mjt} = a_m + \gamma (S_{mj} \times T_t) + \delta_{j,t} + \alpha_2 C_{mjt} + \lambda_2 X^{85}_{m,t} + \eta_{mjt} \tag{7}$$

where Y_{imjt} refers to the log hourly wage of worker i in municipality m, at time t, and T_t is a dummy variable that takes the value of one (zero) for the year 2009 (1998). These cross-sections include the subset of rural municipalities that were sampled in both years of the survey. Since only a few of these towns belong to the same province, I use department-year fixed effects to account for income changes induced by the change in the terms of trade ($\delta_{j,t}$). Hence, I also control for initial urbanization rates ($X^{85}_{m,t}$) to account for differential trends in development across towns of the same department.

In contrast to the analysis in Equations 6, I use a two-sample 2sls approach to estimate the Equations in 7.[37] This approach allows me to estimate the *first-stage* with the full sample of municipalities to improve the precision of these estimates.[38] I follow Pacini and Windmeijer (2016) to correct the standard errors of this two-step procedure and obtain estimates that are robust to heteroskedasticity.

The exclusion restriction in Equation 6 is that municipality's terrain inclination does not lead to different labor market trends across contiguous economies within the same province.[39] This assumption would be violated if the design and execution of public poli-

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

[38]Naturally, the second stage is estimated using only the subset of towns with data on workers' wage.

[39]Formally, $E(s_{mp} \Delta u_{mp} | \Delta C_{mp}, \delta_p) = 0$. This assumption is weaker than the one required for the within

cies in a province depends on topography, or if the trade-induced income effects - associated with the production of certain crops- are not fully accounted for. In Table 4, I show that rural municipalities with distinct levels of terrain inclination had similar trends in urbanization and population growth before trade liberalization.[40] While the exclusion restriction in untestable, these results provide reassuring support for this assumption. In the case of Equation 7, the exclusion restriction is stronger as the larger variation in agro-climatic conditions across municipalities of the same department imply that crop-specific income effects are not fully accounted for . For this reason, the results of the analysis on wages should be regarded with caution.

Conditional on these assumptions, β in Equations 6 and 7 identifies the average marginal effect on the subpopulation of economies that respond to encouragement from the instrument (i.e. compliers).[41] This subpopulation comprises the towns where the shift in land consolidation was driven by the physical and financial ability of production at scale.

4.4 Results

Table 2 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

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. $E(\Delta D_{mp}\Delta u_{mp}|\Delta C_{mp}, \delta_p) = 0$.

[40] These two variables are closely related to changes in sectoral employment and labor supply in a local economy. I will analyze pre-trends on my outcomes of interest once I am granted re-access to the microdata of the 1985 population census at the Colombian National Department of Statistics.

[41] Monotonicity is an additional assumption that is required to interpret β as the local average marginal effect of land consolidation on the outcomes 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 3, 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).[42]

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 A1 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 4), which are measures of urbanization that relate to my outcomes of interest.[43]

In Table 5, 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 2, I also present the Kleibergen-Paap F-statistic and the inference results that provide correct coverage with potentially weak instruments.[44] The standard

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

[43]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.

[44]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 5, 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 large-farm consolidation on rural labor markets in Colombia. The empirical findings indicate that consolidation can lead to a reallocation of labor away from agriculture within rural economies along with a decrease in workers' earnings. These results imply a reduction in local labor demand after consolidation, and are consistent with a model that features differences in farm labor intensity and non-homothetic consumption growth in the nonfarm sector.

My findings on workers' earnings suggest that migration was not large enough to offset the decrease in labor demand in the medium term. This highlights the existence of barriers to geographic labor mobility and aligns with previous work on the delayed response of local labor markets to economy-wide shocks (Dix-Carneiro and Kovak, 2019; Dix-Carneiro, 2014).

The results in this paper have two main implications for the analysis of land policies. First, they shed light on the distributional impacts of large-farm consolidation across individuals as rural workers may experience a decrease in income despite productivity gains. Second, they imply that the reallocation of workers out of the farms can be accompanied by a decline in rural wages. This relationship is the opposite sign than we usually observe with the process of structural transformation and is consistent with recent evidence on the impacts of such transformation across space (Eckert and Peters, 2018).

While this paper is a step toward understanding the effect of large-farm consolidation on rural labor markets, it opens several questions for future research. For example, what is the contribution of each of the mechanisms? And how do the effects vary under different levels of geographic labor mobility? The theoretical framework that I propose can also be used to quantify the impact of the upcoming land reform on 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.

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.

Arteaga, Julián, 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.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, June 2010, *100* (3), 1008–45.

Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song, "Trade Adjustment: Worker-Level Evidence," *The Quarterly Journal of Economics*, 2014, *129* (4), 1799–1860.

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.

_, Gabriel Garber, and Jacopo Ponticelli, "Capital Accumulation and Structural Transformation*," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 1037–1094.

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, *¡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 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.

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, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885.

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

— **and** — , "Margins of labor market adjustment to trade," *Journal of International Economics*, 2019, *117*, 125–142.

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.

Fajgelbaum, Pablo and Stephen J. Redding, "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914," NBER Working Papers 20217, National Bureau of Economic Research, Inc June 2018.

Farrokhi, Farid and Heitor S Pellegrina, "Global Trade and Margins of Productivity in Agriculture," Working Paper 27350, National Bureau of Economic Research June 2020.

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.

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.

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, "La agricultura colombiana en la década del noventa," *Revista de Economía del Rosario*, 1998.

_ , *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.

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.

Laskievic, Andrei Arminio, "Commodity booms and structural transformation: the role of input use and land inequality," WorkingPaper July 2021.

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 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).

Martinez, Luis R., "Sources of Revenue and Government Performance: Evidence from Colombia," WorkingPaper October 2020.

McArthur, John W. and Gordon C. McCord, "Fertilizing growth: Agricultural inputs and their effects in economic development," *Journal of Development Economics*, 2017, *127*, 133–152.

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.

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.

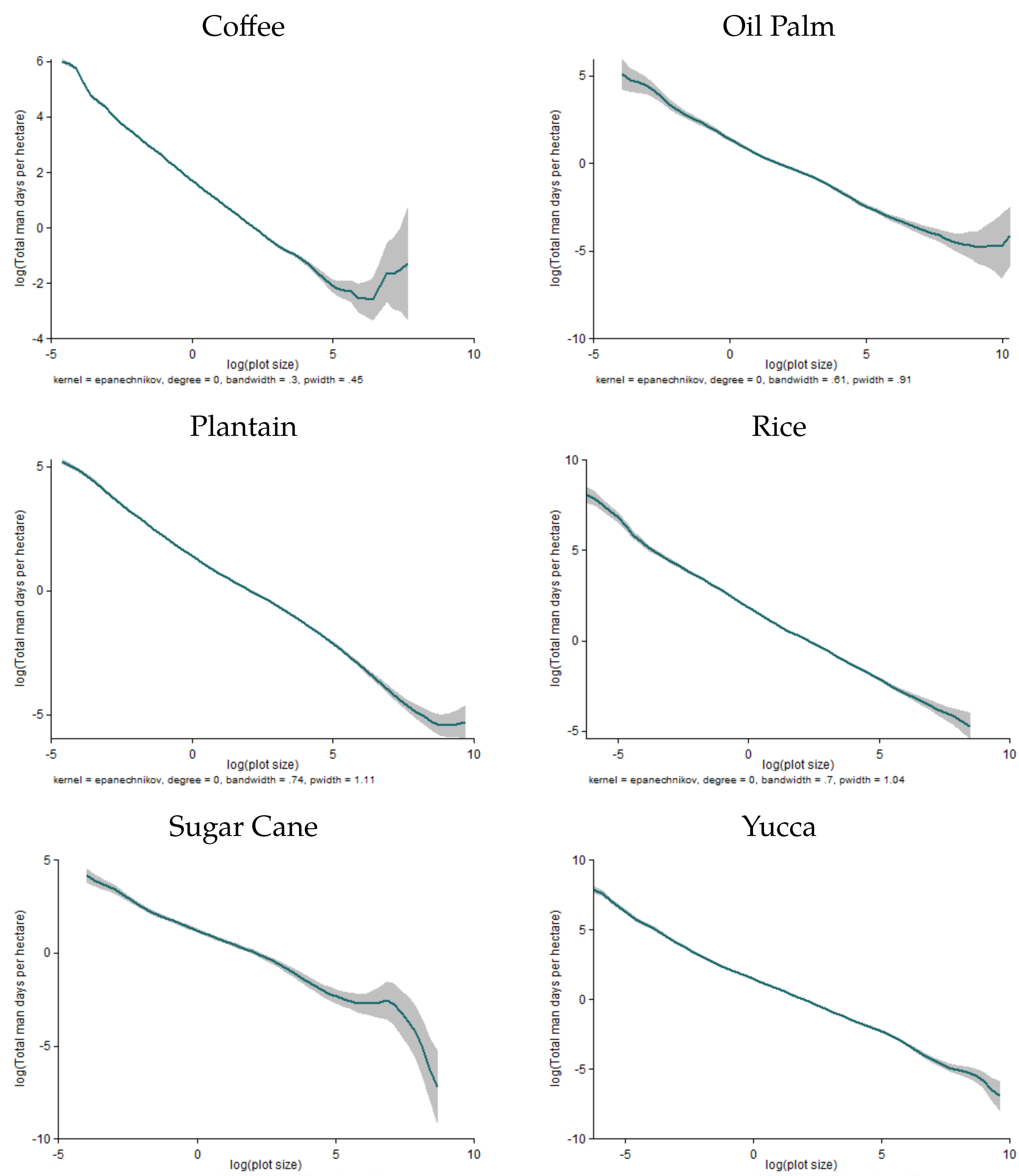
Sotelo, Sebastian, "Domestic Trade Frictions and Agriculture," *Journal of Political Economy*, 2020, *128.*

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

Uribe Castro, Mateo, "Caffeinated Development: Export Sector, Human Capital, and Structural Transformation in Colombia," WorkingPaper February 2020.

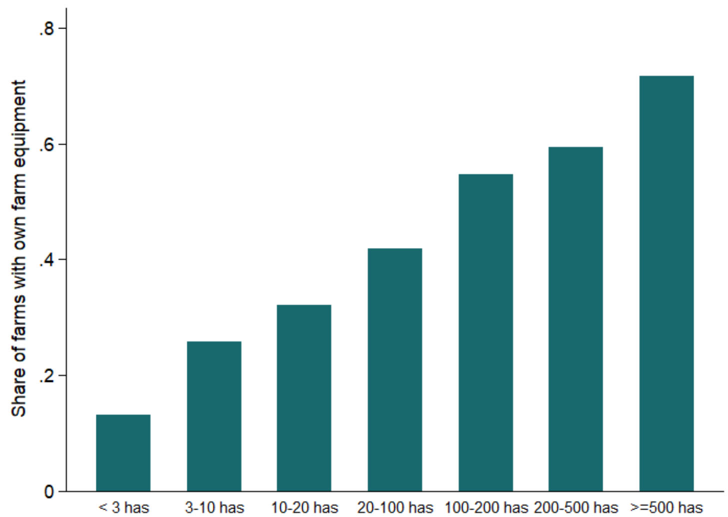
Figures and Tables

Figure 1: Labor Intensity By Plot Size



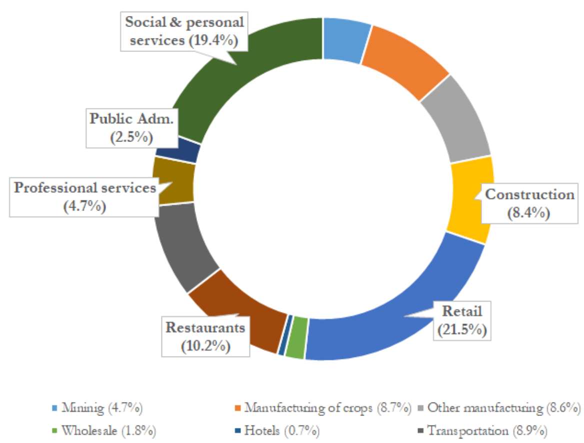
Notes: Own illustrations using data from the 2014 National Agricultural Census.

Figure 2: Equipment Ownership By Farm Size



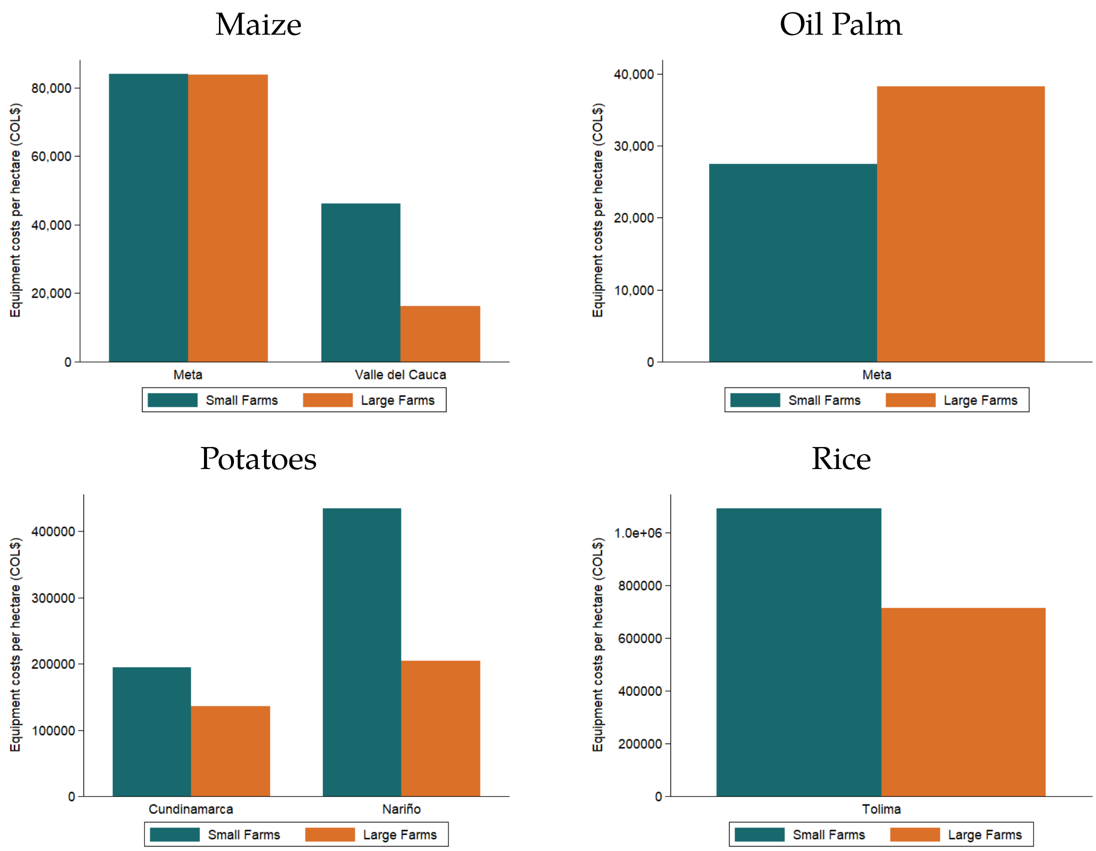
Notes: Own illustrations using data from the 2014 National Agricultural Census.

Figure 3: Distribution of Nonfarm Employment



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

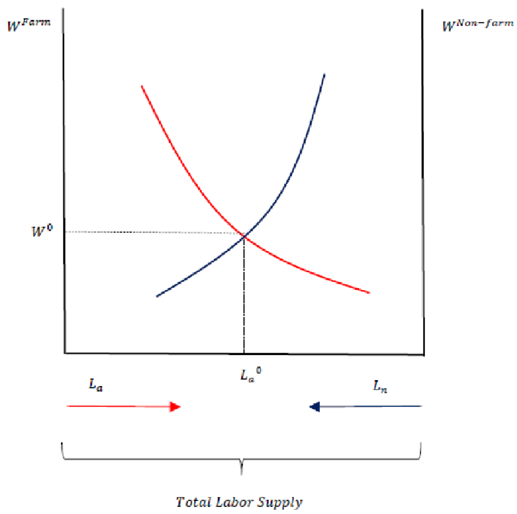
Figure 4: Equipment Costs Per Hectare, By Farm Size



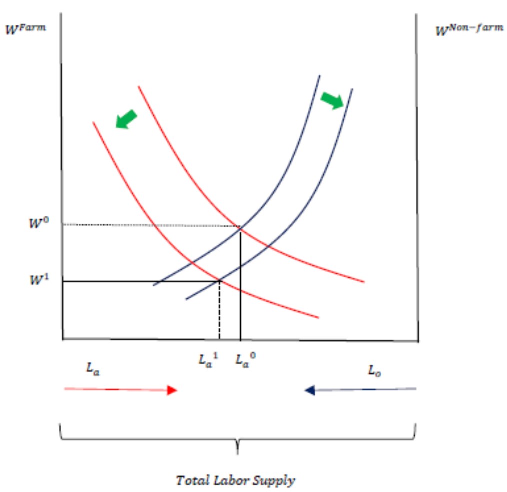
Notes: Own figures based on the data reported in Perfetti et al. (2012). Equipment costs per hectare are measure in Colombian pesos of 2012. Small farms refer to the farms that are at most one family farm unit in size, while large farms refers to farms that are more than one family farm unit in size. For details on the family farm unit measure, see Section 4.1.

Figure 5: Overall Effect on Equilibrium Wage: An Illustration

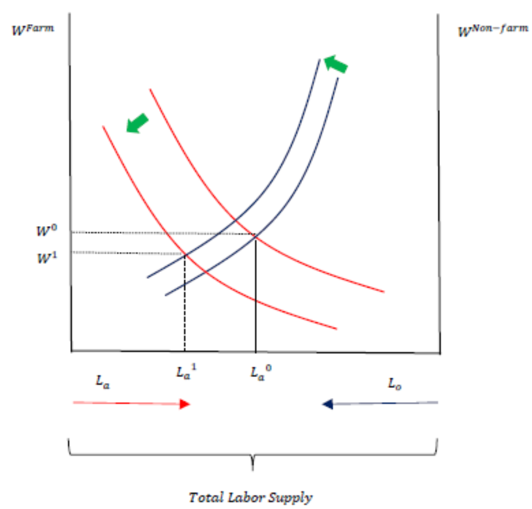
Panel A. The Rural Labor Market



Panel B: Unambiguous Effect

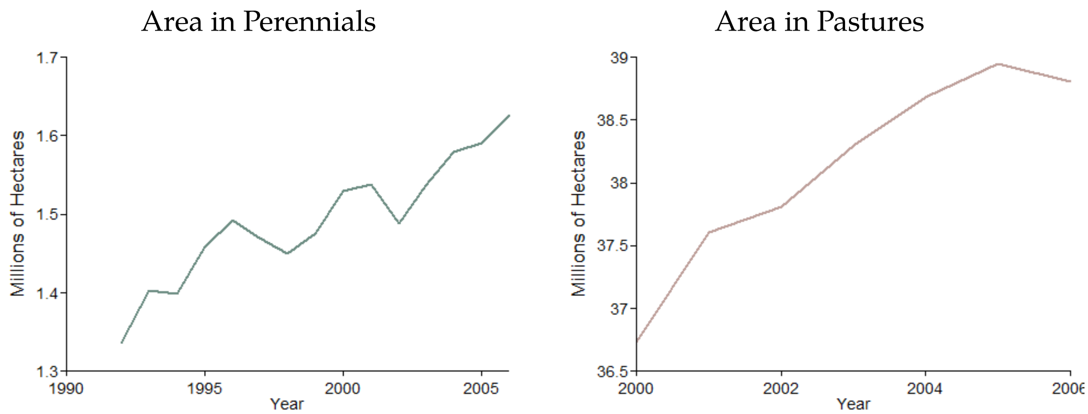


Panel C: Ambiguous Effect



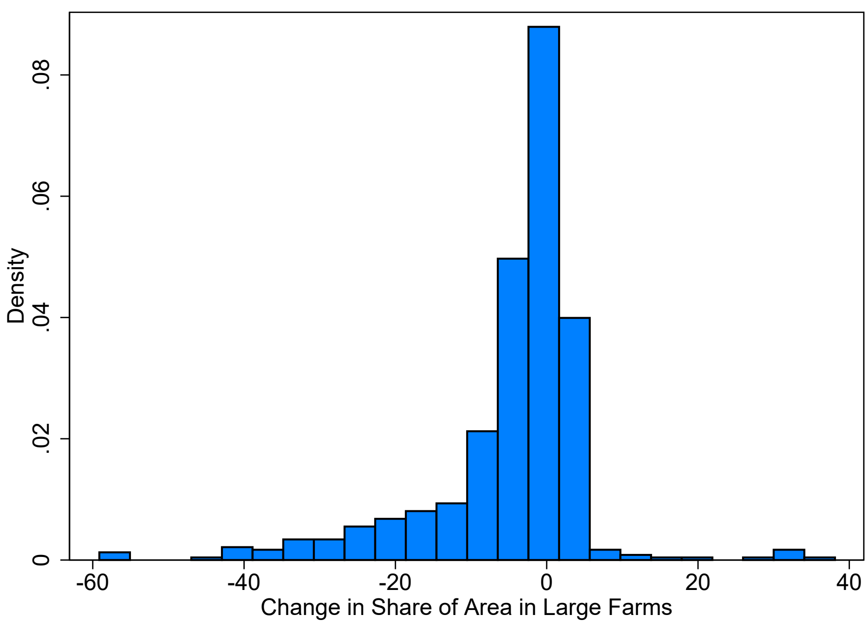
Notes: Own illustration.

Figure 6: Area in Land-Intensive 1993-2005



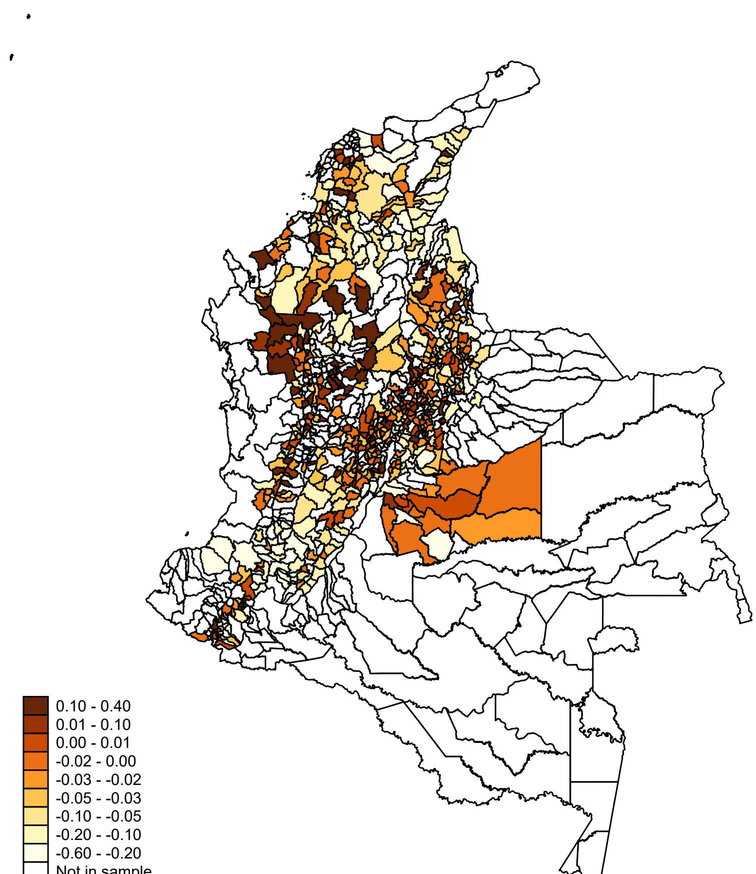
Notes: Own illustrations using data from the Municipal Agricultural Evaluations.

Figure 7: Shift in Large-Farm Consolidation Across Municipalities



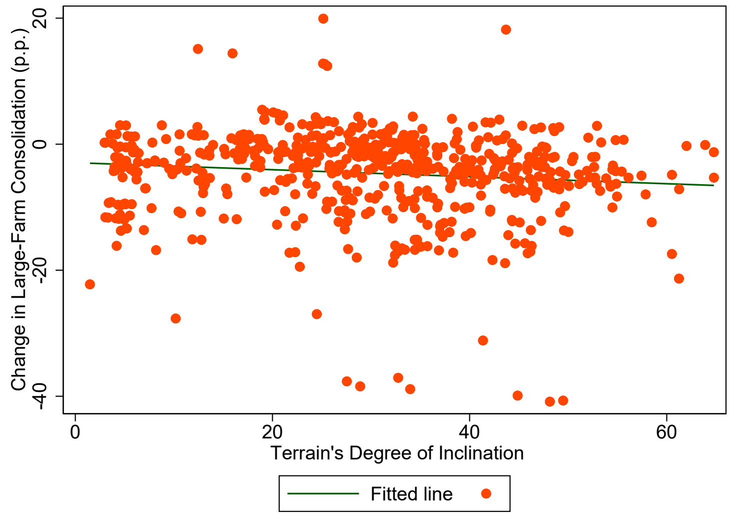
Notes: This figure displays the change in the share of area in large farms (in percentage points) for rural municipalities between 1993 and 2005. Own calculations using cadastral data.

Figure 8: Spatial Variation in Consolidation Shift



Notes: This map displays the change in the share of area in large farms (in percentage points) for rural municipalities between 1993 and 2005. Own calculations using cadastral data.

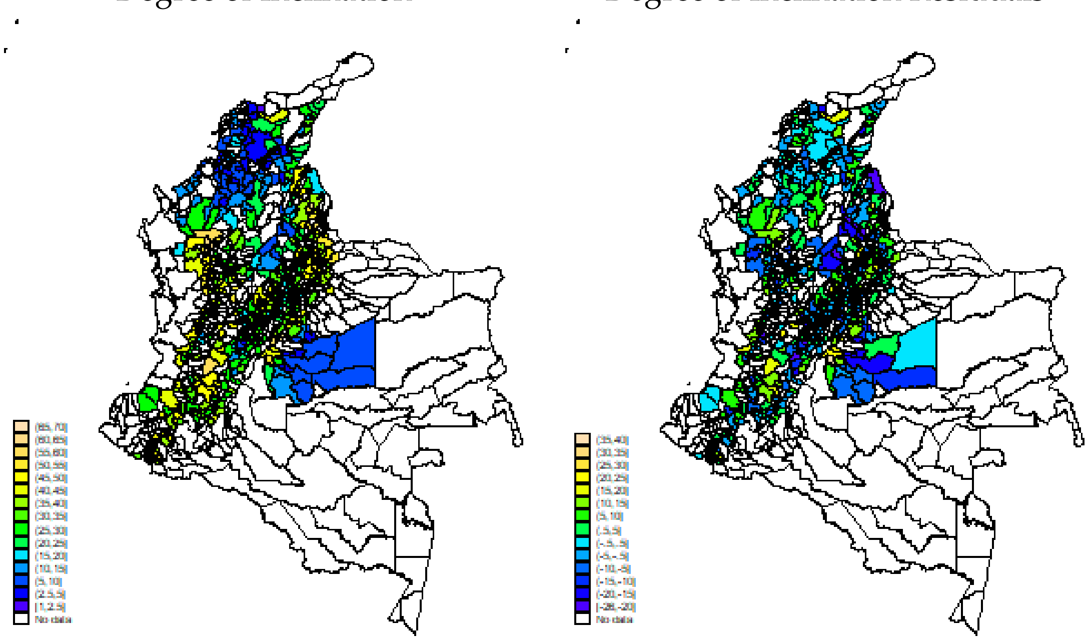
Figure 9: Change in Large-Farm Consolidation & Terrain's Inclination



Notes: This figure depicts the fitted values of a regression of the change in consolidation on the terrain's inclination and province fixed effects.

Figure 10: Spatial Variation in Terrain's Inclination

Degree of Inclination | Degree of Inclination Residuals



Notes: These maps display the spatial distribution of the municipality's terrain inclination (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 and 4.3 for details on the data sources for the analysis on wages.

Table 2: Effect of Large-Scale Consolidation 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 3: Effect of Large-Scale Consolidation 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 4: Large-Scale Consolidation, Sectoral Employment & Unemployment Rate: Pre-Trends Test

	Share Rural Population (Δ_{85-93})		Total Population (Δ_{85-93})	
	(1)	(2)	(1)	(2)
Terrain's Inclination (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 5: Effect of Large-Scale Consolidation 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

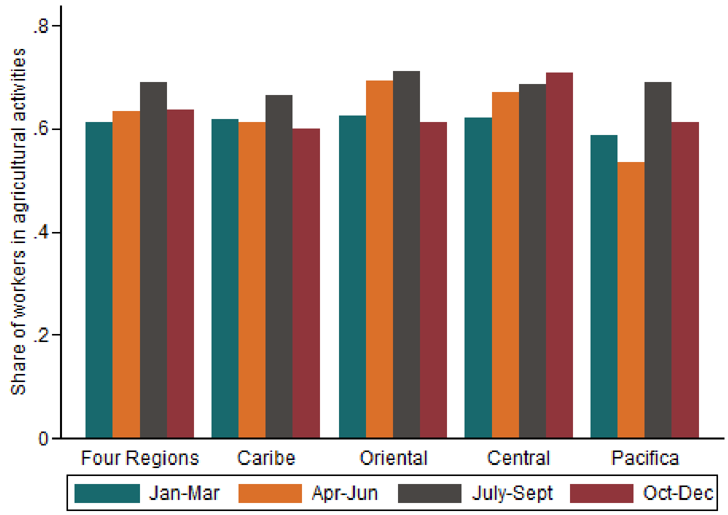
Table 6: Large-Scale Consolidation and Log Hourly Wage: Pre-Trends Test

	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: Large-Scale Consolidation, 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 B2 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).[45] 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 agents is Pareto, which is a standard way to model land heterogeneity in the development literature: $F(h) = 1 - \left(\frac{h_m}{h}\right)^{\tau}$.[46] 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.[47] 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.[48] 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.

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 B3 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

[45] 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.

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

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

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

Table B2: 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.

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.[49] The targeted moments are the aggregate share of final expenditure in agricultural goods and nonagricultural tradable industries.[50]

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.[51] 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.

[49]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: $u(c) = \omega_a log(c_a - \bar{c_a}) + \omega_n log(c_n) + \omega_c log(p_c c_c + p_c\bar{c_c}) - log(p_c)$.

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

[51]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%.

Table B3: 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).