Credit and Saving Constraints in General Equilibrium: A Quantitative Exploration*

Catalina Granda Franz Hamann Cesar E. Tamayo†

Abstract

In this paper we build an incomplete-markets model with heterogeneous households and firms to study the aggregate effects of saving constraints and credit constraints in general equilibrium. We calibrate the model using survey data from Colombia, a developing country in which informal saving and credit frictions are pervasive. Our quantitative results suggest that reducing savings costs increases selection into formal saving, but the effect on aggregate outcomes and welfare is dwarfed by that of a policy which ameliorates borrowing constraints. Such a policy improves resource allocation and increases returns to capital and labor, resulting in higher savings and welfare gains for both households and firms.

Keywords: saving constraints, credit constraints, financial inclusion, misallocation, savings, formal and informal financial markets.

JEL Classification Numbers: E21, E44, G21, O11, O16

*This version has been accepted for publication, after peer review, but is not the version of record and does not reflect post-acceptance improvements, or any corrections. The version of record is available online at: https://doi.org/10.1016/j.jdeveco.2019.06.007.

We thank the Co-editor (David Lagakos) and two anonymous referees for their very detailed suggestions. We also thank conference and seminar participants at the AEA Meetings (Philadelphia, 2018), LACEA-LAMES Meetings (Guayaquil, 2018), EcoMod (Venice, 2018), EEA Meetings (Lisbon, 2017), CEF-SCE (New York, 2017), IFABS (Oxford, 2017), SBIF (Santiago, 2017), ICESI, EAFIT, Banco de la República and Universidad del Rosario for their helpful comments. Amalia Rodríguez provided valuable assistance in microdata processing. All errors and omissions are our own. We gratefully acknowledge funding from the Inter-American Development Bank (IDB, CMF/IFD Division) at the early stages of this project. The views expressed in this paper are entirely those of the authors and no endorsement by the IDB or Banco de la República is expressed or implied.

†Granda: Universidad de Antioquia, catalina.granda@udea.edu.co; Hamann: Banco de la República, fhamansa@banrep.gov.co; Tamayo: Universidad EAFIT, ctamayot@eafit.edu.co

1 Introduction

Financial inclusion has become a priority for development economists and policy makers around the world.[1] In recent years, the longstanding goal of improving access to credit has been joined by a growing interest in the role that saving should have in a comprehensive financial inclusion strategy.

While the literature on credit frictions is well developed and includes both empirical and theoretical contributions, the literature on the causes and consequences of exclusion from formal saving markets (i.e., through financial institutions) mostly comprises field experiments in relatively small communities.[2] In fact, little is known about the general equilibrium effects of saving constraints, or the ways in which they may interact with other frictions, such as those found in credit markets. Our goal is to present a framework that can be used to quantify these effects and to study these interactions.

In this paper, we develop a model of heterogeneous agents in which financial market frictions distort credit and saving decisions by households and firms. In the tradition of Aiyagari (1994) and Huggett (1993), worker-households save for precautionary reasons. Our framework allows them to do so using either a deposit contract (formal saving) or cash (informal saving). Saving constraints result from the fact that using the deposit contract is costly. Entrepreneurs can access credit markets when choosing their capital input, but face collateral requirements due to limited enforcement problems. Saving constraints lead households to seek informal savings instruments (cash) and result in lower aggregate saving, while credit constraints imply firm-level distortions and misallocation of capital across producers as in Restuccia and Rogerson (2008) and Hsieh and Klenow (2009).

To discipline the model parameters, we use data from the Colombian Longitudinal Survey (henceforth ELCA), which contains income and occupational data as well as detailed information concerning financial decisions by households. Importantly, ELCA respondents report the extent to which they use informal saving instruments and why. We complement this information with firm-level and other data from credit markets and macroeconomic aggregates.

Our counterfactual exercises suggest that, in terms of macroeconomic aggregates and welfare, the effects of reducing formal saving costs are dwarfed by the impact of alleviating credit constraints. The intuition for this result is simple: decreasing saving constraints

[1] According to the Alliance for Financial Inclusion (2016), by 2015 over 35 countries had committed to implementing or had already implemented financial inclusion strategies.

[2] Reference studies from the credit frictions literature are Kaplan and Zingales (1997) and Buera et al. (2011), and from the saving constraints literature are Dupas and Robinson (2013) and Karlan et al. (2014).

primarily shifts the aggregate capital stock (by a relatively small amount), while decreasing credit constraints lowers misallocation, which increases aggregate TFP, and thus generates larger welfare gains and more formal savings.

More specifically, due to the fixed nature of saving costs, reducing them has little impact on dynamic decisions. Lowering saving costs increases the fraction of savings allocated to formal instruments, which modestly reduces consumption volatility –by about 2 percent– of those who were already saving (the deposit contract gives them a non-zero return). However, reducing saving constraints leaves the allocation of capital and labor unchanged, and thus has little effect on productivity, wages or welfare.

In turn, ameliorating credit frictions, such that credit to output becomes similar to that of the U.S., allows more productive firms to grow, increasing returns to both capital and labor. Workers welfare rises by 9.4 percent due to an increase in wages, while entrepreneurial welfare is 8.7 percent higher. All in all, TFP and output increase by 2 percent and 10 percent, respectively, while the fraction of income captured by the two bottom quintiles rises from 8 percent to 10 percent.

Higher returns to capital due to rellocation also induce more saving along both intensive and extensive margins: the share of workers who save increases from 30.1 percent to 37.9 percent and the workers saving rate increases from 10.8 percent to 12.6 percent. The fraction of workers who save informally also drops, from 38 percent to 26 percent. Higher saving in this closed economy results in higher investment and the capital intensity rises from 1.74 to 1.97. Moreover, in the economy with lower credit frictions only 51 percent of that capital is financed by entrepreneurs, compared with 82 percent in the benchmark, credit-constrained economy.

To better understand why reducing credit constraints is a much more powerful tool to increase savings and welfare, we conduct a series of partial equilibrium simulations. We first show that changes in saving behavior and welfare from reducing saving constraints are small and identical under general and partial equilibrium settings. This is in sharp contrast with a policy that ameliorates credit frictions, which brings about large welfare gains in either setting. The main difference with the general equilibrium result, is that, in partial equilibrium, benefits are exclusively captured by entrepreneurs. Consequently, income distribution in this case worsens –the fraction of income captured by the two bottom quintiles falls from 8 percent to 6.7 percent. In general equilibrium, welfare gains are more evenly distributed. First, the interest rate adjusts to the higher credit demand, moderating gains in output and entrepreneurial welfare, and increasing returns to saving by workers. Second, and perhaps more importantly, workers capture some of the improvement

in resource allocation from reducing credit frictions via higher wages.

We also provide two extensions to the baseline model and discuss how the main results change. First, we integrate into our setup the fact that lower income households appear to face higher relative costs of using formal savings instruments. Under this framework, reducing saving costs brings about larger increases in formal saving, but has little impact otherwise. Second, we relax the assumption of fixed occupations, allowing for some entry/exit into entrepreneurial activity. Since workers can use bank deposits as collateral for credit in the future, bettering access to credit in this setting becomes an even more powerful tool to foster formal saving. Financial frictions are also less distortive in this case, as agents can choose not to operate a firm under very tight credit constraints.

This paper is related to a number of recent studies addressing the interaction between formal and informal financial markets in developing countries. For example, Wang (2019) develops and estimates a dynamic equilibrium model of borrowing and saving decisions that allows him to interpret Thailand's financial reform in 2001 as one that reduced formal borrowing interest rates, lowered costs of access to credit, and relaxed collateral constraints. This reform in turn led to an increase in the share of households borrowing formally and to a fall in informal interest rates. He finds that the welfare gains from these policies are smaller than those suggested by previous studies that disregard informal saving options.

Furthermore, our paper links to two streams of the financial development literature.[3] One stream has been looking into the determinants of access to and use of savings instruments, and their effect on economic outcomes. This literature has mainly focused on the extensive margin and includes cross-country studies (Demirgüç-Kunt and Klapper, 2013; Rojas-Suarez and Amado, 2014) as well as country-level studies (Beck et al., 2017), and field experiments inside villages or larger regions within a country (see Dupas and Robinson, 2013; Kast and Pomeranz, 2014; Prina, 2015, to name a few).

Overall, this strand of the literature shows that households –particularly those in poor regions– often save using formal or informal instruments that are costly, entail high risk, and have limited functionality. This leads to low saving rates, with significant welfare consequences: reduced consumption smoothing, low resilience to shocks, and foregone profitable investment opportunities. In a survey of this literature, Karlan et al. (2014) group constraints to saving into five categories: transaction costs, lack of trust and regulatory barriers, information and knowledge gaps, social constraints, and behavioral biases. This paper focuses on the first two categories, as these comprise market frictions that hinder

[3]For a recent survey of the financial development literature, see Fernandez and Tamayo (2017).

the supply of savings products.

The other stream of the literature has focused on using structural equilibrium models to quantitatively examine the impact of credit frictions on economic development. In this class of models, financial frictions –typically in the form of collateral constraints– generate distortions in the allocation of capital across productive units that in turn lead to aggregate productivity losses (Buera et al., 2011; Buera and Shin, 2013; Midrigan and Xu, 2014). The size of these losses depends critically on the underlying persistence of productivity shocks, which determines the extent to which entrepreneurs are able to save their way out of collateral constraints (Moll, 2014).

This class of models has successfully been used to study the impact of reducing distortions in credit and factor markets in developing countries. In an influential paper, Kaboski and Townsend (2011) study a large microfinance program in Thailand through the lens of a structural equilibrium model of borrowing constraints, which distort the allocation of capital. The estimated model is able to broadly reproduce the response in credit and consumption that followed the increase in credit availability. More recently, Lagakos et al. (2018) use a dynamic-incomplete markets model to study whether spatial (rural-urban) distortions in labor markets can explain productivity and welfare gaps in developing countries. This literature has also cautioned against the indiscriminate use of market interventions –such as credit subsidies in the Thailand microfinance program– that may have long-term distortionary effects (Fulford, 2013; Antunes et al., 2015; Buera et al., 2013).

Dabla-Norris et al. (2015) analyze three types of financial frictions: participation costs, collateral requirements, and costly monitoring. Their results suggest that the effect of policies alleviating these frictions individually or jointly depends on country-specific characteristics. Using this approach, Karpowicz (2014) finds that, for Colombia, lowering collateral requirements promises higher growth, but that reductions in participation costs offer a better way to reduce inequality.

Missing in the literature are studies quantifying the efficiency gains from ameliorating distortions in the allocation of credit and savings through formal financial instruments. Filling this void is important for at least two reasons. First, development experiments have revealed large general equilibrium effects of relocating toward less risky saving and production technologies (Flory, 2018; Donovan, 2018). Second, a significant determinant of the demand for formal savings instruments is its return, which is an endogenous outcome of the financial intermediation process, and thus, is affected by credit allocation.

The paper is organized as follows. Section 2 presents some empirical regularities pertaining to barriers to financial inclusion and patterns of saving behavior in Colombia. The

main aspects of the model economy are described in Section 3. Section 4 presents simulations and discussion of policy scenarios. Section 5 introduces some extensions to the baseline model. Section 6 concludes.

2 Empirical Regularities

To lay the groundwork for our quantitative model, we build a set of empirical regularities from the Colombian economy using recently collected survey-level data on households' saving habits, as well as firm-level data capturing access to finance. We begin by extracting a number of stylized facts from the Colombian Longitudinal Survey (ELCA), which provides specific information on financial inclusion and the use of financial services for saving.[4]

One striking feature from the ELCA data is the fact that nearly three-quarters of respondents (73 percent) indicated that they do not set aside any fraction of their earnings as savings (see Table 1).[5] This contrasts with predictions from the standard Bewley-Huggett-Aiyagari type of models of non-zero assets holdings, and suggests that saving frictions may be important. Another notable figure that emerges from the survey is the fact that nearly one third of the respondents who were savers in 2013 reported saving "mainly" outside of the financial system (i.e., they use informal financial instruments).[6]

It is not surprising to find that informal saving is more prevalent among the lower income households.[7] The data show that the fraction of savers who do not use the formal financial system is 3.5 times as large in the lowest four income quintiles as in the highest

[4]The ELCA is a household survey recently designed and implemented by the Universidad de los Andes. Data from two waves of the survey (2010 and 2013) have been published so far, and the third wave was rolled out in 2016. For methodological details, see Bernal et al. (2014). Our computations are based on responses by workers from the urban module of the survey, as the rural module lacks representativeness with regard to a number of variables of interest.

[5]Information concerning savings habits is elicited from household members over 10 years of age. Each member is asked whether he/she usually saves part of the income received, and those who respond affirmatively are subsequently asked how much they save each month. For ELCA-based calculations in Table 1 and thereafter, we use data collected from these and other questions in the savings chapter and, in addition, we build earnings data for each individual using the remaining chapters.

[6]The ELCA surveyors ask people where do they "mainly" save. Respondents are given the options of (a) bank or financial institution, (b) cash, (c) employee funds, (d) saving clubs or chains, and (e) other instruments. We assume that all savings in employee funds is channeled through the financial system.

[7]Notice that, while the incidence of informal saving decreases with income, this phenomenon is pervasive even among the highest income deciles. This is in line with evidence from Solo and Manroth (2006), who show that the "unbanked" in Bogota –Colombia's capital city– are not exclusively found in the lower income segments of the population. In this regard, recent studies using ELCA data suggest that labor formality is positively correlated with the likelihood of saving using formal financial instruments (Iregui-Bohórquez et al., 2018). For a model of the relationship between informality and saving behavior through the lens of a dynamic equilibrium model, see Granda and Hamann (2015).

income quintile (59 percent compared to 17 percent). As we shall see, the relationship between income and informal saving is most likely mediated by non-convexities in the cost of using financial instruments (e.g., non-proportional fees), and by geographic and literacy characteristics.

Table 1: Incidence and composition of savings (workers)

	% of respondents		Of those who save	
	do not save	save	% save formally	saving rate (avg)
2010	72.9%	27.1%	61.5%	16.7%
2013	73.3%	26.7%	62.2%	12.1%

Source: Authors' calculations based on ELCA.

According to the ELCA, informal saving appears to be mostly a phenomenon associated with the high costs and low returns of using formal financial instruments for saving. Indeed, Figure 1 (left panel) suggests that taxes, fees and other charges constitute an important motivation for such saving patterns.[8] However, alongside the mentioned costs, low returns to savings also appears to be a crucial factor keeping savers from using the formal financial sector. These ELCA figures are consistent with data from the World Bank's Global Financial Inclusion Indicators (Global Findex); according to these data, Colombia is one of the countries in which a significant number of respondents (about 20 percent) say they do not have an account at a financial institution because they find it too expensive.[9] Given its level of development, Colombia appears to be somewhat of an outlier in this respect (see Figure 1, right panel); concerns about formal banking expenses are higher in Colombia than in some Latin American peers, such as Brazil, Chile, Uruguay, and Costa Rica.[10]

As suggested above, the motivations for using informal savings instruments are also closely related to income levels. Again, using data from the ELCA, Table 2 (left panel) shows that the probability of not saving formally because of high costs or long distances

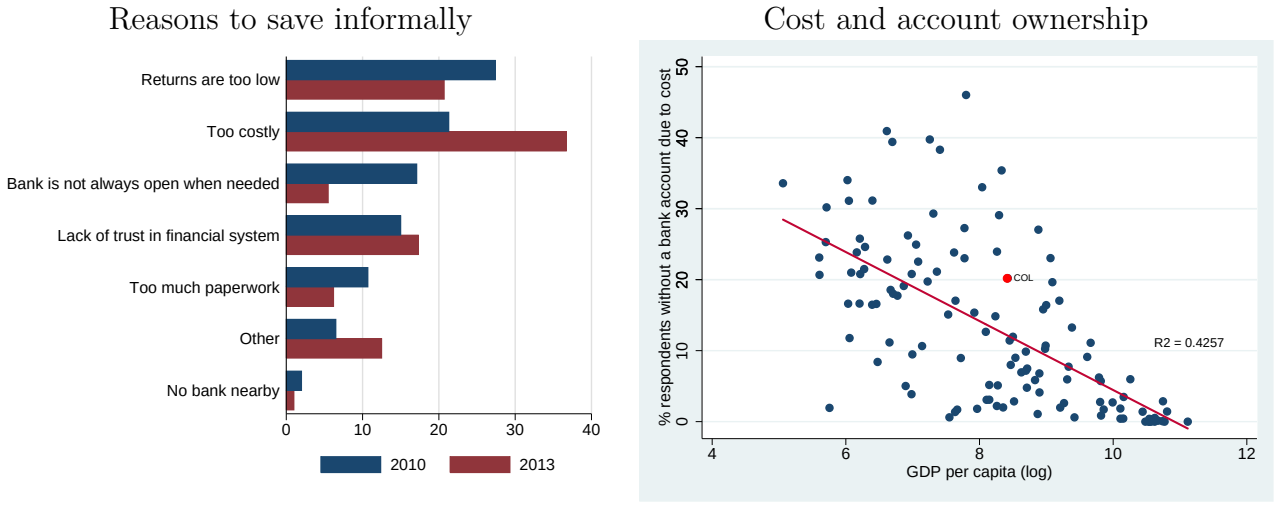
[8]It is worth noting that the main reason adduced for not saving in the 2013 ELCA was "too little money", so we concentrate on the remaining ones. Also, the questionnaire of the 2013 wave included reasons for not saving in the financial system that were not explicitly considered in the 2010 wave and that we grouped in Figure 1 as follows: "too costly" comprises "maintenance fees and use charges are high" and "4 x 1000 is too expensive"; "other" includes "does not know how to", "believes bank would refuse to open account", and "bank refused to open account", among others.

[9]The Global Findex asks respondents if they own an account at a "bank or credit union (or another financial institution, like a cooperative in Latin America)" (Demirgüç-Kunt and Klapper, 2013, p. 313).

[10]A comprehensive volume that documents low –as well as informal– savings in Latin American is Inter-American Development Bank (2016).

to banks decreases with income. Moreover, lack of trust in the financial system –another major reason to save informally– has to some extent been associated with financial literacy, which in turn is closely connected with income levels. Using data from the first nationally representative survey on financial capability, carried out by the Central Bank of Colombia and the World Bank in 2012, Table 2 (right panel) shows that a one-unit increase in monthly income leads to a 0.35 increase in the probability of performing better in the financial literacy quiz associated with the survey.[11]

Figure 1: Reasons to save outside the financial system



Source: Authors calculations based on ELCA and Global Findex.

The ELCA also makes it clear that people in Colombia save mainly for precautionary reasons: In 2013, 30.7 percent of savers reported to have saved for unexpected expenses, while only 15.2 percent and 15.5 percent did so for education and retirement, respectively. These figures are consistent with those of the financial capability survey (see Reddy et al., 2013), which finds that 35.5 percent of Colombia's savers save for unforeseen events.

When making their saving decisions, the precautionary motive is perhaps the only one in which returns to savings may be relatively unimportant. In this case, savers should exhibit a preference for liquid instruments that typically offer very low returns, and that,

[11]The quiz consists of five questions addressing different aspects of financial knowledge and literacy (basic numeracy, time value of money, interest paid on a loan, calculation of principal and interest, and compound interest rate). Based on the number of correct answers, the authors of the study build a financial literacy index taking six possible integer values (0-5) (see the description in Reddy et al., 2013). For our analysis, this index is regressed on (the natural logarithm of) monthly income. Specifically, a generalized Poisson regression model for underdispersed count data is used to take into account the nature of the dependent variable. Marginal effects are reported in the table.

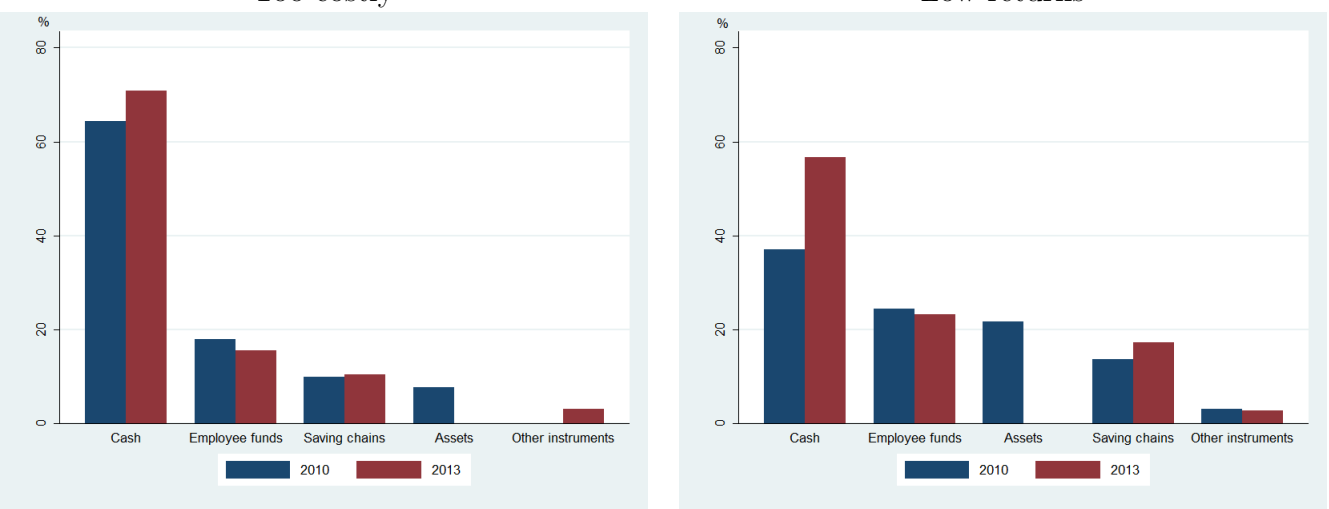
in turn, render transaction costs (fees, taxes, and other charges) as the main determinant in the choice of instrument. It is therefore not surprising that those reporting that the financial system is "too costly" as their reason for not saving through a financial institution, save mostly by holding cash (Figure 2, left panel), while those claiming to save informally because the financial system offers "low returns" save relatively less in cash and more through saving chains (or similar schemes) and other instruments (Figure 2, right panel).

Table 2: Income level and informal saving

	ELCA (logit) save informally because...			Financial Literacy (count data) # of correct answers in quiz	
	too costly	too far away	either	all sample	urban
ln(income)	-0.020	-0.004	-0.027	0.353	0.323
s.e.	(0.007)	(0.001)	(0.007)	(0.032)	(0.048)
Dummy 2013	-0.000	-0.007	-0.009		
s.e.	(0.014)	(0.003)	(0.014)		
Constant	0.470	3.688	1.185	-0.607	-0.395
s.e.	(0.748)	(1.654)	(0.727)	(0.163)	(0.227)
Observations		2,668		1,231	627

Source: Authors calculations based on ELCA and the Financial Capability Survey.
Note: Marginal effects are reported.

Figure 2: Savings instruments according to the reason to save outside the financial system



Source: Authors' calculations based on ELCA.

High costs of service typically result from a combination of fixed costs of infrastructure

and a small scale of operation. One can think of these factors as the underlying reason for the negative relationship between GDP per capita and the widely held view that the financial system is too expensive to use. That is, more developed countries have larger financial systems allowing providers to fully exploit economies of scale.

Low returns on savings are more difficult to rationalize in an economy in which capital is scarce, and inflation is relatively low and predictable, as has been the case in Colombia for the last two decades. In fact, data from the International Monetary Fund show that real interest rates on money market instruments during the 2000-2014 period were lower in Colombia than in Latin American peers such as Peru, Mexico, Uruguay, and Brazil. Such low returns on savings can be explained either by costly intermediation or low returns on investment. In this latter respect, our calculations using the methodology of Caselli and Feyrer (2007) suggest that in recent decades the marginal productivity of aggregate capital has been lower in Colombia than in other Latin American countries such as Mexico, Chile, and Peru, as well as in emerging Asia.[12] What can lie behind such low returns to investment? Recent studies suggest that financial frictions distorting credit allocation are usually associated with this phenomenon. Such frictions make capital flow toward less profitable projects, which in turn results in aggregate productivity losses and a lower economy-wide return to capital (see Buera et al., 2011; Midrigan and Xu, 2014).

Table 3: Firms' financial constraints and saving patterns

Enterprise Survey (2010)		Firm-level data (average 2010-2013)	
Firms reporting financial constraint	41.5%	Firms not saving	26.6%
Firms financing investment with debt	35.0%	Saving rate (savers)	9.4%
% of investment financed by debt	21.2%	Leverage (A/E)	2.46

Source: World Bank and Paez and Tamayo (2019).

One natural place to look for potential capital misallocation issues lies in indicators of financial constraints. In this respect, data from the World Bank Enterprise Survey (WBES) reveal that 41.5 percent of Colombian firms reported access to finance as being a major obstacle to their operations in 2010. This places Colombia as the fourth country

[12] Aggregate marginal productivity of capital (MPK) is defined as $MPK_R = (1-\theta)\alpha\,(Y/K_R)\,(P_y/P_{kR})$, where P_y/P_{kR} is the inverse of the relative price of capital obtained as the ratio of the investment deflator to the GDP deflator using data from the World Development Indicators; $1-\theta$ is the ratio of reproducible capital to total capital (including natural resources) obtained from the World Bank database "Wealth of Nations"; and α, K_R, and Y are the share of capital in output, reproducible capital, and GDP, respectively, all taken from the Penn World Tables. Using data for 2005, this methodology yields an MPK of 7.6 percent in Colombia, whereas it is 8.4 percent in a typical Latin American country, 9.3 percent in a typical country in emerging Asia, and 11.1 percent in a typical advanced economy.

in Latin America in which firms find themselves more credit constrained. Moreover, only 35 percent of these firms reported having funded investment by borrowing from banks (see Table 3, left panel).

Limited access to external financing usually forces firms to accumulate net worth in order to invest and post collateral. In Colombia, firm-level data collected by Paez and Tamayo (2019) suggest that this is precisely the case: In a sample of 231,222 firms between 2010 and 2013, 73 percent accumulated net worth at the considerable rate of 9 percent to 10 percent of total operating income (see Table 3, right panel).[13] Another indication that firms in Colombia may have limited access to external finance is their relatively low leverage, as measured by the assets-to-equity ratio; the ratio for Colombian firms is 2.46, compared to ratios of 4.5 to 5 for European firms, as recently reported by Kalemli-Ozcan et al. (2012).

3 A Model of Credit and Saving Constraints

In this section, we develop a dynamic, heterogeneous agents, general equilibrium model in the spirit of Aiyagari (1994), Huggett (1993) and Restuccia and Rogerson (2008). The model features saving constraints at the household level, and credit constraints at the firm level. Some of the noteworthy features of the model come directly from the evidence found in the ELCA and the World Bank databases: saving in banks is costly; people save mainly for consumption smoothing; the main alternative to bank saving is cash; and firms face borrowing constraints.

The economy is populated by a measure N of workers and a unit measure of entrepreneurs. Both workers and entrepreneurs are heterogeneous with respect to their productivity and seek to maximize lifetime utility given by

$$\mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t) \right\},$$

where period utility is of the constant relative risk aversion form:

$$u(c) = \frac{c^{1-\chi}}{1-\chi},$$

with $\chi > 0$, and $\beta \in (0,1)$ being the discount factor.

Entrepreneurs can borrow and save with financial intermediaries, but they face a col-

[13] See Appendix A.1 for a brief summary of the dataset and its sources.

lateral requirement that constrains the amount they can borrow. Workers face uninsurable idiosyncratic labor income risk and have access to financial markets. There are two types of financial instruments available: a one-period risk-free asset (formal) and cash (informal). Notice that the income risk and our assumption about preferences force workers to save for consumption smoothing. That is, more and better saving opportunities should allow for lower consumption volatility and thus higher welfare.

Workers. Each worker is endowed with a unit of labor that is supplied inelastically. Labor income, however, depends upon the worker's idiosyncratic efficiency, ε_t, which is random, and evolves over time according to a finite-state Markov process with transition probabilities $\psi(\varepsilon', \varepsilon) = Pr(\varepsilon_{t+1}|\varepsilon_t)$ and ergodic distribution $\Psi(\varepsilon)$.

Workers save using cash, s, or a one-period deposit contract, q. Those who engage in deposit savings must pay a fixed cost, $\xi = \tau w$, for every period they use the deposit contract, where w is the wage and $\tau \geq 0$ is a parameter. With this fixed cost of using "formal" financial instruments, we aim to capture the fact documented in Section 2 that a sizable fraction of savers do not use the financial system because they find it too costly.[14] Particularly, our specification of the cost as a multiple of the wage can be thought of as reflecting fees and commissions charged by financial institutions. These charges are often intended to cover the personnel costs of running bank branches, and to provide certain services required for deposit account management.[15] In Section 5.1 below, we extend this interpretation to cover other non-pecuniary costs of saving associated with distance and/or financial literacy. Note that while cash –the "informal" instrument– yields no interest, deposits –the "formal" instrument– yield a non-negative risk-free rate of return, r.

Given prices (r, w), a worker's problem can be stated recursively as

$$W(q, s, \varepsilon) = \max_{c,q',s'} \frac{c^{1-\chi}}{1-\chi} + \beta \sum_{\varepsilon'} W(q', s', \varepsilon')\psi(\varepsilon'|\varepsilon) \tag{1}$$

subject to

$$c + q' + s' + \xi \mathbb{I}_{q'>0} = w \exp(\varepsilon) + (1 + r)q + s, \tag{2}$$

[14]It is possible that one reason why people do not save formally to create a buffer stock is that they are relatively well insured though informal means. This would imply that the value for τ calibrated below in Section 4.1 could be too high relative to its true value.

[15]In this vein, recent evidence by Roa and Carvallo (2018) shows that monthly fees for maintaining a deposit account in Colombia can be as high as 1.5 percent of the median monthly wage. It is worth noting that this study updates and extends the exercise conducted in Beck et al. (2008), which first collected data on the costs of owning and using financial services around the world.

and the no-borrowing constraints

$$q \geq 0,\ s \geq 0,$$

where $\mathbb{I}_{q'>0}$ is an indicator variable that equals one if the worker saves using the formal instrument, and zero otherwise.

Entrepreneurs. Entrepreneurs have access to a decreasing returns technology that uses labor l and capital k to produce a consumption good, y. Specifically,

$$y_t = [\exp(z_t)]^{1-\theta}(k_t^{\lambda} l_t^{1-\lambda})^{\theta}. \tag{3}$$

In (3) $1-\lambda \in (0,1)$ governs the share of labor in production, while $\theta < 1$ is the degree of decreasing returns to variable inputs. Capital depreciates between periods at rate δ. An entrepreneur's idiosyncratic productivity is given by z_t, which evolves over time according to a finite-state Markov process with transition probabilities $\pi(z', z) = Pr(z_{t+1}|z_t)$ and ergodic distribution $\Pi(z)$. In each period, a fraction $1-\eta$ of entrepreneurs dies and is replaced by new ones, in which case the firms they owned exit at zero market value.[16]

Entrepreneurs decide how much to borrow (d_t) and save (b_{t+1}). Since $b_t = k_t - d_t$, and b_t is pre-determined, choosing d_t amounts to choosing k_t. Further, we assume that entrepreneurs cannot fully commit to repaying loans because financial contracts are imperfectly enforceable. In particular, defaulting entrepreneurs keep a fraction $1-\phi$ of their capital stock; the remaining fraction ϕ is recovered by the lender.[17] Finally, we assume that all entrepreneurial saving is made through the one-period deposit contract.[18]

Given prices, (r, w), an entrepreneur's problem can be stated recursively as

$$V(b, z) = \max_{b', k, l} \frac{c^{1-\chi}}{1-\chi} + \beta\eta \sum_{z'} V(b', z')\pi(z'|z) \tag{4}$$

subject to

$$c + b' + \xi = [\exp(z)]^{1-\theta}(k^{\lambda} l^{1-\lambda})^{\theta} - (r+\delta)k - wl + (1+r)b$$

[16] This reflects that firms exit the market for reasons not internalized by the model. It is well known that without exogenous exit some firms would eventually accumulate enough assets to overcome borrowing constraints, so that over time the mass of firms would grow without bound (Quadrini, 2004).

[17] Note that although the collateral constraint ensures that all contracts are enforceable, if a firm were to default, it would exit the market irreversibly.

[18] This is in line with the evidence. According to Didier and Schmukler (2014), virtually all firms in Latin America own and use formal bank accounts. For Colombia, in particular, the WBES data show that between 96 percent and 99 percent of the surveyed firms own checkings or savings accounts.

and the collateral constraint

$$d \leq \phi k,$$

which can be rewritten as

$$k \leq \frac{b}{1-\phi}.$$

Notice that our specification of the collateral requirement is virtually identical to that used in Midrigan and Xu (2014).

Financial intermediaries. Banks take deposits from workers and lend them to entrepreneurs. Because all contracts are strictly enforceable (i.e., there is no default in equilibrium), entrepreneurs pay and workers receive exactly the risk-free rate, which is endogenously determined. Naturally, some firms face a higher shadow price of capital than others depending on whether the collateral constraint binds, and some workers face a lower return once fixed costs of deposit market participation are accounted for.

Equilibrium. The economy has a stationary equilibrium that consists of a set of prices (w, r), stationary distributions of workers, $g(\cdot)$, and entrepreneurs, $h(\cdot)$, and decision rules $\{c(q,\varepsilon), q'(q,\varepsilon), s'(q,\varepsilon), b'(b,z), k(b,z), l(b,z)\}$, such that:

- All workers and entrepreneurs optimize, that is, $l(b,z), k(b,z), b'(b,z)$ solve problem (4)-(7) and $c(q,\varepsilon), q'(q,\varepsilon), s'(q,\varepsilon)$ solve (8)-(9);
- The labor market clears,

$$\sum_{b,z} h(b,z) l(b,z) = N \sum_{\varepsilon} \varepsilon \Psi(\varepsilon);$$

- The asset market clears,

$$\sum_{b,z} h(b,z) k'(b,z) = \sum_{q,s,\varepsilon} g(q,s,\varepsilon) q'(q,s,\varepsilon) + \sum_{b,z} h(b,z) b'(b,z).$$

4 Quantitative Performance

In this section, we describe how data are used to calibrate the model presented in Section 3. We also present a series of policy experiments that allow us to quantify the costs associated with saving and credit constraints in a developing economy such as that of Colombia.

4.1 Calibration

The model is calibrated to be consistent with a number of features of the Colombian economy. In order to better match the data, we add an aggregate efficiency component, A_t, which grows deterministically at a constant rate g, $A_t = gA_{t-1}$. This implies that most aggregates in this economy are non-stationary with a deterministic trend. Normalizing $A_0 = 1$ and defining $\gamma = g^{1/(1-\alpha)}$ with $\alpha = \lambda\theta$, such trend becomes γ^t.[19]

We divide the parameter vector into two groups. The first group includes preference and technology parameters that are difficult to identify using our data (see Table 4). We assign to these parameters values that are common in the existing dynamic general equilibrium (DGE) literature. Accordingly, the period is set to one year and the effective discount factor is equal to 0.951. This is within the range of values commonly found in studies of emerging market economies. Likewise, the risk aversion coefficient is set to 2.3, which is close to the value estimated for Colombia in Prada and Rojas (2010).

Table 4: Preference and technology parameters

Parameter	Value	Description	Source
$\beta(1-\gamma)^{1-\chi}$	0.951	Effective discount factor	DGE literature
χ	2.300	Risk aversion coefficient	DGE literature
θ	0.850	Share of variable inputs	Zuleta et al. (2010)
$1-\lambda$	0.635	Labor share in variable inputs	Zuleta et al. (2010)
δ	0.075	Capital depreciation rate	Hamann et al. (2013)
$1-\eta$	0.07	Firm exit rate	Eslava et al. (2013)
γ	1.038	Trend output growth	DANE

As for the technology parameters, Zuleta et al. (2010) apply several methodologies to estimate the factor shares during the 1984-2005 period. We take averages of some of their obtained series for the shares of basic labor, physical and human capital, such that the share of labor $1-\lambda$ in variable inputs is set to 0.635, and the degree of decreasing returns θ sums to 0.85. These figures are virtually identical to those found in Atkeson and Kehoe (2005) and Restuccia and Rogerson (2008) for the United States, and similar to figures that have previously been used for Colombia (Granda and Hamann, 2015).

Also, the depreciation rate δ is set to 0.075 as in Hamann et al. (2013). Further, the survival rate of entrepreneurs η is set so that $1-\eta = 0.07$ to match the average firm exit

[19] Trend growth also implies that the effective discount rate is $\beta(1-\gamma)^{1-\chi}$ for workers and $\beta\eta(1-\gamma)^{1-\chi}$ for entrepreneurs. The stationary equilibrium of the model of course is obtained in terms of de-trended variables. For further details, see Appendix A.2.

rate in the manufacturing sector as reported in Eslava et al. (2013). Finally, the trend growth parameter γ corresponds to the long-run output growth rate, and is estimated as the average annual growth rate of output from 1976 to 2012 using yearly data from the national statistics office (DANE).

Values for the second group of parameters are chosen to replicate certain moments of the Colombian data (see Table 5). The transitory productivity of workers and entrepreneurs,(ε, z), are assumed to be first-order autoregressive processes with Gaussian disturbances, and are discretized into five-state Markov chains using the Rouwenhorst (1995) method. The autocorrelation coefficients, $(\rho_\varepsilon, \rho_z)$, and the standard deviations, $(\sigma_\varepsilon, \sigma_z)$, are chosen to approximately match the saving rate and the fraction of non-savers for workers and entrepreneurs, respectively. The former is obtained from the financial module of the 2013 wave of the ELCA, while the latter is computed as the ratio of changes in net worth to total operating revenue from the firm-level dataset (also for the year 2013) of Paez and Tamayo (2019).

Table 5: Summary of calibrated parameters

Param	Value	Description	Target	Source
ρ_ε	0.65	AR(1) labor productivity	% of workers who do not save	ELCA
σ_ε	0.24	Std dev labor productivity	Workers' saving rate	ELCA
ρ_z	0.28	AR(1) entrep productivity	% of entreps who do not save	ELCA
σ_z	0.76	Std dev entrep productivity	Entrepreneurs' saving rate	ELCA
ϕ	0.20	% of pledgeable collateral	Credit-to-output ratio	Central Bank
τ	0.08	Fixed cost of formal saving	% of formal savers	ELCA

Finally, we calibrate the parameters that govern the functioning of financial markets. The cost of using formal savings instruments τ is set to match the fraction of savers that resort to formal financial instruments in the ELCA data. Similarly, the parameter that captures limited enforcement ϕ is chosen to replicate the ratio of credit to enterprises (corporate plus microcredit) to private value added computed using data from Banco de la República –the central bank of Colombia.

The resulting economy, as can be seen in Table 6, resembles the targeted moments fairly well. Specifically, the model economy appropriately replicates the fraction of non-savers, the saving rates, the fraction of households that save using formal financial instruments, and the credit-to-output ratio.

Table 6: Calibration results

Targeted moment	Data	Model
% of workers who do not save	73.3%	69.6%
% of formal savers	62.2%	62.1%
Workers' saving rate	12.1%	10.8%
% of entrepreneurs who do not save	28.8%	31.1%
Entrepreneurs' saving rate	9.9%	8.5%
Credit-to-output ratio	31.8%	31.6%

4.2 Counterfactual Analysis: Baseline Results

We now use the calibrated model to analyze a number of financial inclusion policies. In addition to looking at key macroeconomic aggregates and distributional statistics, we study the welfare implications of alternative policies by measuring conditional welfare changes, often called consumption equivalent variations. The idea, first introduced by Lucas (1987) and extended to heterogeneous agents models by İmrohoroğlu (1989), is to measure how much consumption of an agent needs to change in every state in the stationary equilibrium so that the agent would be indifferent between experiencing the effects of the policy and living in the pre-reform economy. If we denote $\{c_t^B\}_{t=0}^{\infty}$ as the optimal consumption plan in the benchmark economy (in this case calibrated to Colombia) and $\{c_t^P\}_{t=0}^{\infty}$ as the plan resulting from implementing a given policy, the welfare metric for workers in our model is a function $\omega(q, s, \varepsilon)$ such that

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u((1 + \omega(q, s, \varepsilon))c_t^B) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t^P).$$

For our choice of reward function $u(\cdot)$, this computation is given by

$$\omega(q, s, \varepsilon) = \left[\frac{W^P(q, s, \varepsilon)}{W^B(q, s, \varepsilon)}\right]^{1-\chi} - 1,$$

where $W^P(q, s, \varepsilon)$ and $W^B(q, s, \varepsilon)$ solve (1) for the worker (q, s, ε) in the policy and benchmark economy, respectively. Identical calculations are made for entrepreneurs. In what follows, we present results from the average welfare gain using the stationary distribution from the pre-reform (i.e., benchmark) economy (e.g., $\bar{\omega}^W = \sum_{q,s,\varepsilon} g(q, s, \varepsilon)\omega(q, s, \varepsilon)$ for workers), as well as the distributions of such gains by income quintile.

The first experiment aims to measure the impact of reducing costs associated with

formal savings, τ. In principle, it is not straightforward to choose a lower value for this parameter because microdata from ELCA-type surveys are not readily available for other countries. To get around this issue, we use data from the Global Findex to obtain a rough approximation to the fraction of people in the United States that save using mostly informal instruments.[20] We then lower τ up to the point where we match this fraction. The approximation obtained is 13.5 percent, which corresponds to $\tau = 0.045$. The results from such a reduction are presented in the third column of Table 7, wherein, to facilitate comparison, we reproduce the performance of the benchmark economy under the label "model Colombia".

It can be seen that lowering the cost of formal saving entails some benefits for workers, as the fraction of formal savers rises from 62.1 percent to 87.6 percent and their overall welfare increases modestly by 0.6 percent. Welfare gains occur for two reasons. First, some workers who have low savings in absolute value, and who were more likely to save informally, now use the deposit contract. This increases their return to saving and helps them reduce consumption volatility (see bottom-left panel of Figure 3). Secondly, workers who were already saving formally –higher income workers for the most part– experience an increase in average consumption because of the reduction in a cost they were paying.

Entrepreneurial welfare, by contrast, exhibits a significant increase, even though consumption is as smooth as in the benchmark economy. This is because entrepreneurs now experience a fall in a cost they are forced to pay. Notice that these gains are concentrated among entrepreneurs in the bottom quintiles (see Figure 3) as they see a larger proportional increase in their consumption (recall saving costs do not depend on income or savings). All other aggregates including output, total factor productivity and saving rates remain unchanged.

The next policy we consider aims to alleviate credit frictions while keeping saving constraints in place. In particular, we maintain the cost of saving formally at its calibrated value ($\tau = 0.08$) but, instead, loosen the collateral constraint by increasing ϕ up to the point where the business lending to output ratio resembles that of the U.S. economy (97

[20] In particular, we look at data from the Findex and compare it with data from the ELCA. While the latter asks individuals where they mainly saved, respondents to the Findex survey can report all the instruments they used for saving, meaning that the fraction of respondents saving informally includes those who did only a tiny portion of their savings that way. For Colombia, the Findex reports that around 73 percent of respondents used cash and other informal instruments to do any of their savings, which is 92 percent higher than the figure from the ELCA, so we apply this factor to the Findex data for the U.S. Thus, we divide 26 percent –the number reported as people saving informally by the Findex– by 1.92 to get an approximation to the fraction of U.S. people who *mainly* save outside the financial system.

percent).[21] This corresponds to $\phi = 0.54$, and the results are presented in the fourth column of Table 7. The most noteworthy finding from this experiment is that reducing credit frictions is a more powerful tool to increase workers saving and overall welfare. This results from both a higher return to capital –i.e., bank deposits; notice the increase in formal saving– and a higher labor income due to better resource allocation –the wage rate is 11 percent higher.

Table 7: Policy experiments: Baseline Results

Statistic	Model "Colombia"	τ = US ϕ = COL	τ = COL ϕ = US	τ = US ϕ = US
Targeted moments				
% of workers who do not save	69.6%	70.0%	62.1%	56.2%
% of formal savers (workers)	62.1%	87.6%	73.9%	95.0%
Workers' saving rate (savers)	10.8%	10.6%	12.6%	16.7%
% of entrepreneurs who do not save	31.1%	31.1%	35.1%	35.1%
Entrepreneurs saving rate (savers)	8.5%	8.5%	7.2%	7.2%
Credit-to-output ratio	31.6%	31.7%	96.3%	96.8%
% income in quintiles 1 & 2	8.0%	8.0%	9.8%	10.0%
Non-targeted moments				
% of capital financed by firms	81.8%	81.8%	51.1%	50.9%
Capital intensity (K/Y)	1.74	1.74	1.97	1.97
Aggregate output	1.00	1.00	1.10	1.10
Total factor productivity (TFP)	1.00	1.00	1.02	1.02
Net real interest rate	8.1%	8.1%	8.4%	8.3%
Real wage rate	0.37	0.37	0.41	0.41
Consumption volatility				
Workers	0.43	0.42	0.41	0.40
Entrepreneurs	0.49	0.49	0.47	0.46
Welfare gain ($\bar{\omega}$)				
Workers)		0.6%	9.4%	10.1%
Entrepreneurs)		9.2%	8.7%	15.6%

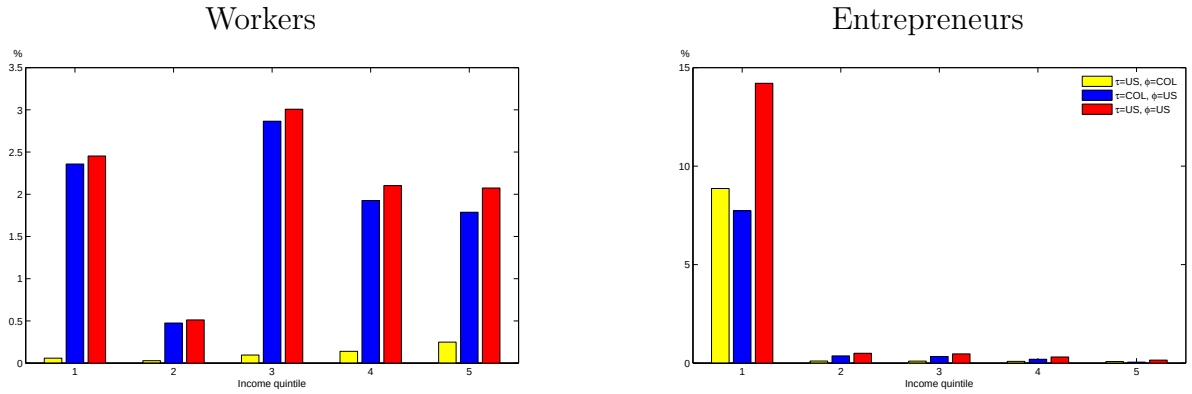
Aggregate output rises by 10 percent and average welfare increases for both workers

[21] While total credit to the private non-financial sector as a share of GDP in the U.S. was 149 percent in 2013, almost half of it was mortgage and consumer credit. The credit to output ratio used here is obtained by adding the bank credit to non-financial businesses as a share of GDP (67 percent) from Table F4.1 of the Bank for International Settlements (https://stats.bis.org/statx/srs/table/f4.1) and the debt securities to GDP ratio (27 percent), also obtained from the BIS.

(9.4 percent) and entrepreneurs (8.7 percent), largely a result of higher incomes.[22] This is especially the case for workers, as a larger fraction of them now save and their saving rate increases to almost 13 percent. As shown in Figure 3, both workers and entrepreneurs in the lowest income quintiles benefit most from this policy, and income distribution becomes less concentrated: the share of income held by quintiles 1 and 2 increases from 8 percent to 9.8 percent. Notice also that only relaxing borrowing constraints makes entrepreneurs save less (lower saving rate). This is because a much higher fraction of the capital stock is financed by household savings (higher credit-to-output ratio).

Figure 3: Welfare Gains by Income Quintile

Consumption Equivalent Variation



Note: The figure displays the distribution of welfare gains by income quintile associated with the counterfactual experiments of Table 7.

In our last experiment, shown in the rightmost column of Table 7, we combine the two policies considered above; that is, we set $\phi = 0.54$ and $\tau = 0.045$. This policy combination has the largest impact on formal saving and on workers saving rates, which climb to 95 percent and 16.7 percent, respectively. The intuition for this result is as follows: When only formal saving costs are lowered, saving does not increase much because demand for credit changes little (recall the closed-economy assumption). At the other end, when only credit frictions are reduced, saving increases because of higher demand for credit and higher interest rates, but saving does not grow as much because the cost of using banks remains in place. When both frictions are reduced simultaneously, workers saving increases

[22] There is also a reduction in consumption volatility, especially among low income workers (see Figure 5 in the Appendix). However, such improvements in consumption smoothing are relatively less important, since the increase in output and productivity allows agents to move toward a flatter part of their utility function, where they care less about volatility.

substantially. Notice, however, that with respect to the credit frictions only reform, this combination of policies has a modest (positive) impact on the remaining aggregates and distributional statistics.

Here we have used figures from the U.S. economy to set up our policy experiments mainly to facilitate comparison with the extant literature. Additional counterfactuals (presented in Appendix A.3) that mimic the much higher financial market development of a country like Sweden suggest that the gains from financial reforms that loosen credit and saving constraints can be much more substantial.[23]

To summarize, the main counterfactual exercises from our baseline model economy show that the benefits of reducing savings costs alone are very modest when compared with those of alleviating credit constraints. This result occurs because fixed saving costs do little to modify dynamic behavior, while alleviating borrowing constraints brings about a powerful general equilibrium effect. Relaxing credit frictions improves resource allocation, which in turn (i) increases the aggregate productivity of capital, inducing higher saving (along both the extensive and intensive margins) and less informal saving, and (ii) increases the productivity of labor, allowing for higher average consumption by workers. The relative contribution of each of these forces in terms of macroeconomic aggregates and welfare is more easily seen by studying the case of a small open economy in which r is fixed, and by providing some partial equilibrium results that remove the effect of higher wages.

4.3 Counterfactual Analysis: Open Economy

The results from the previous subsection suggest that the equilibrium effects of removing saving and credit constraints are mediated by changes in both interest rates and wages. But Colombia has at times been thought of as a rather small open economy (SOE) with domestic interest rates virtually pegged to the foreign ones (López et al., 2008). In this subsection, we provide results from similar counterfactual exercises when the interest rate in the model economy is assumed to remain unchanged at its benchmark level of 8.1 percent.

Table 8 reproduces the model calibration, and presents the results from the counterfactual experiments under the small open economy assumption. In particular, rather than "calibrating" the values for (τ, ϕ) to match the U.S. moments, we use the parameter values used for Table 7, and study their differential impacts.

[23] In particular, workers saving rate increases by 74 percent (8 percentage points from 10.8 percent to 18.8 percent), output increases by 19 percent, and the share of income held by the bottom 40 percent increases by 62 percent (from 8 percent to 13 percent).

It is clear from Table 8 that in a SOE, reducing saving costs has the exact same effect as in the closed economy, but reducing credit frictions has no effect on inducing higher or more formal saving by workers. This is because the interest rate effect is absent: workers do not observe an increase in returns to capital, and, thus, they do not change their saving behavior. However, the welfare gains from alleviating collateral constraints remain because the wage does increase as a consequence of better capital allocation. In our next exercise we isolate this labor market effect as well.

Table 8: Policy Experiments: Small Open Economy

Statistic	Model "Colombia"	τ = US ϕ= COL	τ = COL ϕ= US	τ = US ϕ= US
% of workers who do not save	69.6%	69.8%	70.1%	67.1%
% of formal savers (workers)	62.1%	87.7%	61.0%	84.1%
Workers' saving rate	10.8%	10.7%	9.8%	12.0%
% of entrepreneurs who do not save	31.1%	31.1%	35.0%	35.0%
Entrepreneurs' saving rate	8.5%	8.5%	7.3%	7.3%
Credit-to-output ratio	31.6%	31.6%	98.2%	98.2%
% income in quintiles 1 & 2	8.0%	8.0%	9.6%	9.5%
Real wage rate	0.37	0.37	0.41	0.41
Consumption volatility				
Households (workers)	0.43	0.42	0.41	0.40
Firms (entrepreneurs)	0.49	0.49	0.47	0.46
Welfare gain ($\bar{\omega}$)				
Households (workers)		0.6%	9.8%	10.4%
Firms (entrepreneurs)		9.2%	8.7%	15.8%

4.4 Counterfactual Analysis: Partial Equilibrium

We continue to dissect our baseline results of Table 7 by isolating changes in factor markets. That is, we now conduct counterfactual experiments in an economy in which both the interest rate and the wage remain constant at their benchmark levels ($r = 8.1\%$ and $w = 0.37$). We call this setting, wherein both households and firms solve their problems given prices, partial equilibrium. Results are presented in Table 9 below.

Table 9: Policy Experiments: Partial Equilibrium

Statistic	Model "Colombia"	τ = US ϕ= COL	τ = COL ϕ= US	τ = US ϕ= US
% of workers who do not save	69.6%	69.8%	69.6%	69.8%
% of formal savers (workers)	62.1%	87.7%	62.1%	87.7%
Workers' saving rate	10.8%	10.7%	10.8%	10.7%
% of entrepreneurs who do not save	31.1%	31.1%	35.3%	35.2%
Entrepreneurs' saving rate	8.5%	8.5%	7.4%	7.3%
Credit-to-output ratio	31.6%	31.6%	98.3%	98.3%
% income in quintiles 1 & 2	8.0%	8.0%	6.7%	6.7%
Aggregate output	1.00	1.00	1.51	1.51
Total factor productivity (TFP)	1.00	1.00	1.07	1.07
Consumption volatility				
Households (workers)	0.43	0.42	0.43	0.42
Firms (entrepreneurs)	0.49	0.49	0.48	0.48
Welfare gain ($\bar{\omega}$)				
Households (workers)		0.6%	0.0%	0.6%
Firms (entrepreneurs)		9.2%	30.0%	35.8%

Like the SOE case, reducing saving costs has the exact same result as in the baseline (Table 7). As before, alleviating credit frictions has little impact on workers saving because the interest rate effect is absent. That is, workers will not change their saving behavior unless the explicit cost of accessing formal financial instruments declines. Finally, and perhaps more importantly, notice that when borrowing constraints are relaxed, workers welfare is now unchanged because the wage effect is now absent.

The remaining differences between Table 9 and 7 are fairly intuitive. When collateral requirements are reduced, entrepreneurs can borrow and run larger firms, but they can now hire additional workers at a constant wage (i.e., there is no labor market effect) which allows them to run even larger firms than in the baseline economy. The result is an increase in output of over 50 percent. However, these gains accrue only to entrepreneurs because workers face a constant income (labor supply, interest rates and wages are all fixed). This is why entrepreneurial welfare increases by over 30 percent, and income becomes more (rather than less) concentrated as a consequence of reforms.

5 Extensions

We now present two important extensions to our baseline model in an effort to add some more realism to the analysis. The first extension is concerned with taking seriously the evident inequality in formal saving costs briefly discussed in Section 2. The second extension relaxes the assumption of fixed occupations, allowing for some entry/exit into entrepreneurial activity.

5.1 Saving Constraints: More Heterogeneity

In the model of Section 3, a fixed cost is equally paid by all agents who use the deposit contract. While the fixed nature of this cost induces some inequality on the severity of the saving constraints faced by agents with different income and saving needs, the resulting inequality in the calibrated economy appears to be lower than that observed in Colombia. In this regard, note that, according to data from the ELCA survey, the share of informal savers among workers in income quintiles 1 through 4 is 3.5 times as high as that of workers in the top income quintile; in contrast, the model of Section 3 delivers a ratio of only 2. That across-the-board fees and commissions are not able to replicate this inequality suggests that certain agent-specific characteristics, such as geographic location or financial literacy, may be important determinants of informal saving behavior.

In view of the above, we propose a simple extension to our benchmark model that better captures heterogeneity in saving constraints by making use of the previously established facts that income is negatively associated with the probability of reporting informal saving because of costs or distance, and positively associated with the level of financial literacy (see Table 2 in Section 2). In particular, for workers we posit that saving costs are correlated with income in the fashion:

$$\tau(\varepsilon) = \zeta + \frac{\kappa}{\varepsilon^2},$$

where (ζ, κ) are parameters to be calibrated and ε is the transitory income shock experienced by workers. For entrepreneurs, we continue to assume that they always pay the formal saving cost at a rate $\tau(\bar{\varepsilon})$, where $\bar{\varepsilon}$ is the highest level of the transitory shock (i.e., entrepreneurs pay the lowest formal saving cost).

With this simple modification, we recalibrate our model to replicate the Colombian economy. The only difference with respect to the procedure described in Section 4.1 is that we now calibrate (ζ, κ) to approximately match the fraction of informal savers and the ratio of informal savers in the bottom four income quintiles to that of the top quintile.

The calibrated values for these parameters are $(\zeta, \kappa) = (0.0168, 0.0682)$.

In Table 10, we compare the results from this calibration with the data and with the calibration using the baseline model (without heterogeneous costs of formal saving). We also present the results from three policy experiments similar to those shown in Table 7. Note that when lowering the costs of saving formally, we decrease κ up to the point at which we replicate the informal saving incidence of the United States. The resulting value is $\kappa = 0.0189$.[24]

Table 10: Policy experiments: Heterogeneous cost for formal saving

Statistic	Colombia	Model calibration		Counterfactuals		
				τ=US	τ=COL	τ=US
	Data	τ=for all	$\tau(\varepsilon)$	ϕ=COL	ϕ=US	ϕ=US
Informal saving (workers)	37.8%	37.9%	35.9%	12.1%	24.5%	8.2%
Within income q1-q4	58.0%	55.1%	68.5%	33.2%	52.6%	16.3%
Within income q5	17.0%	27.8%	20.9%	0	0	0
Ratio of q1-q4 to q5	3.5	1.98	3.3	–	–	–
Welfare gain ($\bar{\omega}$)						
Households (workers)				0.7%	10.0%	11.0%
Firms (entrepreneurs)				0.4%	2.1%	2.3%

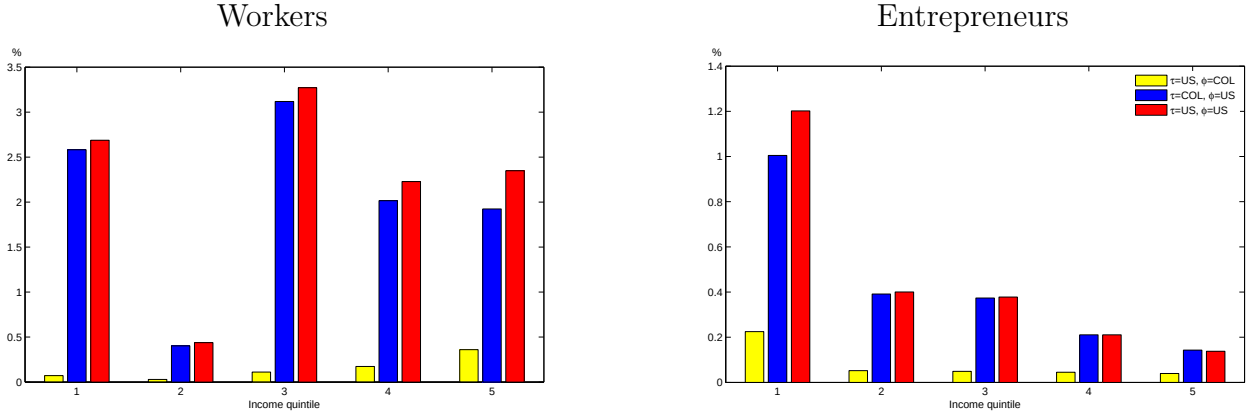
With respect to the baseline model, the model with heterogeneous saving costs is only slightly less effective in matching the overall fraction of informal savers (among workers); however, it comes much closer to matching the actual inequality in informal saving (i.e., the ratio of informal savers in the lower four income quintiles to informal savers in the top quintile). Though it overpredicts the fraction of informal savers in the lower quintiles, it approximates informal saving behavior among the top income quintile (17 percent in the data) better than the baseline model.

The counterfactual experiments reveal that lowering the average cost of formal saving as well as its inequality between income groups results in an increase in formal saving that is larger than that obtained with the baseline model (refer to the last column of Table 7). The increased effectiveness of financial inclusion policies brings about larger welfare gains than in the baseline exercises, especially for workers in the lowest income quintiles (see Figure 4). Other macroeconomic aggregates and distributional statistics experience only marginal changes, and, thus, are not shown in the interest of brevity.

[24] If we were to lower ζ, instead, most results would be identical except for the distribution of welfare gains. In that case, the share of welfare gains that accrues to the lower income households is lower.

Figure 4: Heterogeneous τ: Welfare Gains by Income Quintile

Consumption Equivalent Variation



Note: The figure displays the distribution of welfare gains by income quintile associated with the counterfactual experiments of Table 10.

5.2 Occupational choice

Our model of Section 3 features two separate categories of agents, entrepreneurs and workers, each of them in fixed supply. That is to say, the occupational decision is exogenous. This assumption may seem at odds with recent studies developing quantitative frameworks to address the consequences of entrepreneurship for issues such as wealth inequality and economic development.[25] And while data from the ELCA suggest that saving to undertake entrepreneurial activities may not be very prevalent –only between 4 percent (2013 wave) and 6 percent (2010 wave) of the surveyed individuals report having saved to start a business– it is still useful to study whether formal saving and the possibility of credit that it brings may be an important determinant of occupational choice. For these reasons, in this section we provide a simple occupational choice version of our model as an extension. We describe its main features here, and provide details of the setup in Appendix A.5.

The first modification with respect to the baseline model is that, since agents are now of a single type (i.e., they are not exogenously chosen to be workers or entrepreneurs), there is a single process that captures transitory ability shocks. As before, idiosyncratic efficiency, x_t, evolves over time according to a finite-state Markov process with transition probabilities $\mu(x', x) = Pr(x_{t+1}|x_t)$ and ergodic distribution $\Gamma(x)$.

Without loss of generality, we assume that the occupation choice is made one period in

[25] See Quadrini (2009) and references therein. Also, De Nardi and Fella (2017) tackle the distributional implications of entrepreneurship in a comprehensive survey.

advance. That is, an agent wakes up to every period either as a worker or an entrepreneur. In the former case, the agent inelastically supplies her unit of labor to an entrepreneur, and receives her labor income, $w\exp(x)$, while in the latter case, the agent decides how much labor and capital to use in production.

Agents also decide how much to accumulate in assets. In the case of a worker, she may hold cash, s, or deposits, a, by paying the fixed cost of using the bank, ξ, as defined in Section 3. Likewise, an entrepreneur can only hold assets through the deposit contract, a, but his deposits may be used as collateral to install capital in line with the borrowing constraint $k \leq a(1-\phi)^{-1}$.

Finally, agents choose which occupation they would like to undertake in the following period. For the case of an agent who wakes up to a worker in the current period, the recursive formulation of the problem can be stated as:

$$W(a,s,x) = \max_{c,a',s'} \frac{c^{1-\chi}}{1-\chi} + \beta \sum_{x'} \mathcal{O}(a',s',x')\mu(x',x), \tag{5}$$

where $\mathcal{O}(a',s',x') = \max\{W(a',s',x'), V(a',s',x')\}$, and $W(a',s',x')$ and $V(a',s',x')$ are the continuation values of being a worker and an entrepreneur, respectively. The maximization in (5) is subject to the same resource constraint (2), with appropriate changes of notation (a in lieu of q and x in lieu of ε). The problem of an agent who wakes up as an entrepreneur in the current period can be obtained by making similar adjustments to the problem in (4), and by adding a fixed cost, f_e, that is paid in every period in which the firm operates.[26]

For the quantitative analysis, our strategy is to maintain as much as possible the calibration used in Section 3, and so, we proceed as follows: First, the transitory ability process is parametrized to mimic that of *workers* in Section 3 (i.e., $\rho = 0.65$ and $\sigma = 0.24$). Then, we re-calibrate ϕ and τ so as to once again match the credit-to-output ratio and the fraction of informal savers among workers in Colombia. This results in $(\phi, \tau) = (0.146, 0.059)$. The value for the fixed cost, f_e, is chosen so that the fraction of workers is close to 80 percent, which is a reasonable value for Colombia.[27] Finally, we leave all the remaining parameters $(\beta, \delta, \gamma, \theta, \lambda, \chi, \eta)$ as in the baseline calibration. The results from this exercise are presented in the second column of Table 11.

[26] This is standard in the firm dynamics literature and very much in the spirit of Hopenhayn (1992).

[27] There is considerable uncertainty surrounding this figure for the case of Colombia. While in the ELCA around 61 percent of respondents were classified as workers, Mejia (2010) estimate this number to be closer to 95 percent using the national household survey.

Naturally, we do not expect to match the moments regarding saving behavior as accurately as in the previous settings, since we have only one income process for both workers and entrepreneurs (i.e., two parameters and four moments). Nevertheless, the model can be calibrated to very closely match credit to output and informal saving in the Colombian economy.

We again conduct the counterfactual experiment in which ϕ is lowered until the credit to output ratio matches that of the U.S. economy. This yields $\phi = 0.451$, and the results are displayed in the third column of Table 11. It is worth noting that from this policy scenario, reducing credit frictions becomes an even more powerful tool to reduce informal saving (compare with the fourth column of Table 7). This is not surprising since better access to credit for entrepreneurs incentivizes formal saving by workers, who may become entrepreneurs and use their formal savings as collateral in the future.

Table 11: Policy Experiments: Occupational Choice Model

Statistic	τ = COL ϕ= COL	τ = COL ϕ= US
% of workers who do not save	70.9%	62.5%
% of formal savers (workers)	61.3%	76.9%
Workers saving rate	9.6%	12.2%
Credit-to-output ratio	31.9%	95.7%
Real output	1.00	1.05
Real wage rate	0.53	0.54
Fraction of workers	79.7%	82.6%
Welfare gain ($\bar{\omega}$)		1.1%

Notice also that, as credit frictions diminish, a smaller fraction of agents decides to become entrepreneurs (the fraction of workers is now 83 percent, compared with 80 percent in the pre-reform economy). This is a standard result in the literature (see, e.g., Antunes et al., 2008) and follows from the fact that, with better credit allocation, more productive and larger firms operate, which in turn implies that fewer firms are required to clear the labor market. Finally, note that output and welfare gains from the financial reform are now somewhat lower because, in the pre-reform economy, agents can choose not to operate a firm under very tight credit constraints; hence, financial frictions are less distortive.

6 Concluding Remarks

In this paper, we used recently collected survey data to study the costs associated with saving and credit constraints through the lens of an otherwise standard general equilibrium, heterogeneous agent model. In our setup, the costs of using financial instruments distort saving decisions by households, leading to volatile consumption profiles. These constraints interact with credit frictions to generate a vicious circle of informal savings, capital misallocation, and low returns to formal savings instruments.

Overall, our quantitative results show that alleviating enforcement frictions in credit markets can substantially change saving behavior, and bring about large welfare gains. By contrast, policies that exclusively aim at reducing fixed saving costs can increase the use of formal saving instruments, but have little impact in terms of macroeconomic aggregates and welfare. The intuition for these results is fairly simple: lowering saving costs mostly affects a static choice between saving instruments, while relaxing borrowing constraints improves resource allocation, which, in turn, increases returns both to saving and to labor income. Because these effects are captured only in a general equilibrium framework, our results suggest that this type of study could greatly complement the growing literature on small-scale field experiments associated with financial inclusion policies.

One limitation of the framework developed in this paper is the assumption concerning the means that both formal and informal saving can take. In reality, individuals can resort to financial instruments other than cash and/or a deposit contract; for example, one can save in land, housing improvements, livestock, jewelry, among others. Moreover, in many developing countries, households can put their money into informal savings instruments that might be risky or whose functionality is limited. Future studies should more realistically model these admittedly important features of saving in the developing world.

Another issue to acknowledge is that the process of pricing financial products is a complex one. Banks usually price their deposit products through a combination of fees and interest rates. Some fees may be visible and charged on the "base product" –the deposit account itself, while some others may be what the specialized literature calls aftermarket fees (overdraft fees, ATM fees, etc.). And while the bulk of the industrial organization literature has focused on the explicit interest rate as the main pricing variable, economists have long recognized that set-up costs, fees, commissions and surcharges are important determinants of the demand for deposit accounts (see, e.g., Flannery, 1982; Klein, 1974).[28]

[28] An additional cost of saving formally that we do not consider explicitly in our model is that associated with switching banks. The pervasiveness of network effects in the bank deposit market has been shown to

The available data are widely supportive of the observation that fees and fixed costs on the ownership and use of financial products are a significant source of bank revenue, and therefore constitute a considerable source of price variation. By 2015, in the United States total service fee revenues for banks reached $34.6 billion – approximately 15 percent of total banking non-interest revenue or 5 percent of total bank revenue (Adams, 2017). In the developing world, the incidence of service fees is even more prominent. According to the Global Findex, in Latin America and Sub-Saharan Africa, the overwhelming majority of people who do not have an account at a financial institution report that the main reasons they do not use such institutions are that they are "too costly" (46 percent) and "too far" (65 percent). In Sub-Saharan Africa, lack of documentation also appears to be an important issue. In Latin America, a recent survey of 71 banks in 10 countries shows that deposit accounts are more difficult and costly to open and maintain in this region (Roa and Carvallo, 2018).[29] Finally, the importance of pecuniary and non-pecuniary costs in saving decisions is supported by the literature on field experiments (mentioned above), in which physical proximity and non-fee treatments have resulted in greater account ownership and larger savings (Prina, 2015; Kast and Pomeranz, 2014).

A number of features about the size and structure of the financial system may lie behind the incidence of high service costs to formal saving. To begin with, high fixed costs associated with the provision of deposit services –such as branches and ATMs– along with a small scale of operation (small customer base) results in high unit service costs for financial institutions, which in turn may pass-through to customers. Similarly, low competition among depository institutions can lead to high overhead costs and low efficiency, both of which also translate into high service fees and surcharges. Indeed, according to World Bank data, over the 2010-2015 period, overhead costs as a fraction of assets appear to be higher in Colombia than in the typical Latin American country. Moreover, recent competition estimates by Gomez et al. (2018) suggest that market power by Colombian banks has increased during the same period. Finally, "know your customer" regulations and provisions to stop money laundering and terrorism financing may have also contributed to the persistence of non-pecuniary costs of owning and using formal financial products, especially savings and deposit instruments.

result in considerable switching costs (see, e.g., Kim et al., 2003).

[29]For instance, the survey finds that while in Spain one can open deposit accounts in up to five different channels (branches, correspondents, telephone, etc.), in most Latin American countries this can be made only in two channels. Likewise, in 9 of the 10 countries positive minimum balances are required to open and maintain such accounts, and in 7 of them deposit accounts are charged on average one monthly fee.

References

Adams, R. M. (2017). Bank Fees, Aftermarkets, and Consumer Behavior. Finance and Economics Discussion Series 2017-054, Board of Governors of the Federal Reserve System.

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics 109*(3), 659–684.

Alliance for Financial Inclusion (2016). The 2016 Maya Declaration Report: Celebrating Five Years of Advancing Financial Inclusion.

Antunes, A., T. Cavalcanti, and A. Villamil (2008). The effect of financial repression and enforcement on entrepreneurship and development. *Journal of Monetary Economics 55*(2), 278–297.

Antunes, A., T. Cavalcanti, and A. Villamil (2015). The effects of credit subsidies on development. *Economic Theory 58*(1), 1–30.

Atkeson, A. and P. Kehoe (2005). Modeling and measuring organization capital. *Journal of Political Economy 113 (5)*, 1026–1053.

Beck, T., A. Demirgüç-Kunt, and M. S. Martinez-Peria (2008). Banking services for everyone? Barriers to bank access and use around the world. *World Bank Economic Review 22*(3), 397–430.

Beck, T., H. Pamuk, and B. R. Uras (2017). Entrepreneurial Saving Practices and Reinvestment: Theory and Evidence. *Review of Development Economics 21*(4), 1205–1228.

Bernal, R., X. Cadena, A. Camacho, J. C. Cardenas, L. Fergusson, A. M. Ibáñez, X. Peña, and C. Rodriguez (2014). Encuesta Longitudinal Colombiana de la Universidad de los Andes - ELCA 2013. Documentos CEDE 2014-42, Universidad de los Andes.

Buera, F., B. Moll, and Y. Shin (2013). Well-Intended Policies. *Review of Economic Dynamics 16*(1), 216–230.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review 101*(5), 1964–2002.

Buera, F. J. and Y. Shin (2013). Financial Frictions and the Persistence of History: A Quantitative Exploration. *Journal of Political Economy 121*(2), 221–272.

Caselli, F. and J. Feyrer (2007). The Marginal Product of Capital. *Quarterly Journal of Economics 122*(2), 535–568.

Dabla-Norris, E., Y. Ji, R. Townsend, and D. F. Unsal (2015). Identifying Constraints to Financial Inclusion and Their Impact on GDP and Inequality: A Structural Framework for Policy. IMF Working Paper 15/22, International Monetary Fund.

De Nardi, M. and G. Fella (2017). Saving and wealth inequality. *Review of Economic Dynamics 26*, 280–300.

Demirgüç-Kunt, A. and L. Klapper (2013). Measuring financial inclusion: Explaining variation in use of financial services across and within countries. *Brookings Papers on Economic Activity 46*(1), 279–340.

Didier, T. and S. L. Schmukler (2014). *Emerging Issues in Financial Development: Lessons from Latin America*. The World Bank.

Donovan, K. (2018). Agricultural risk, intermediate inputs, and cross-country productivity differences. Unpublished manuscript, Yale School of Management.

Dupas, P. and J. Robinson (2013). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics 5*(1), 163–92.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2013). Trade and Market Selection: Evidence from Manufacturing Plants in Colombia. *Review of Economic Dynamics 16*(1), 135–158.

Fernandez, A. and C. Tamayo (2017). From institutions to financial development and growth: what are the links? *Journal of Economic Surveys 31*(1), 17–57.

Flannery, M. (1982). Retail bank deposits as quasi-fixed factors of production. *American Economic Review 72*(3), 527–36.

Flory, J. (2018). Formal finance and informal safety nets of the poor: Evidence from a savings field experiment. *Journal of Development Economics 135*, 517–533.

Fulford, S. L. (2013). The effects of financial development in the short and long run: Theory and evidence from India. *Journal of Development Economics 104*(C), 56–72.

Gomez, J. E., C. E. Tamayo, and O. Valencia (2018). Bank market power and firm finance: Evidence from bank and loan level data. IDB Working Paper, Inter-American Development Bank.

Granda, C. and F. Hamann (2015). Informality, Saving and Wealth Inequality in Colombia. IDB Working Paper IDB-WP-575, Inter-American Development Bank.

Hamann, F., I. Lozano, and L. F. Mejía (2013). Sobre el impacto macroeconómico de los beneficios tributarios al capital. In L. E. Arango and F. Hamann (Eds.), *El mercado de trabajo en Colombia: hechos, tendencias e instituciones*, pp. 253–284. Banco de la República.

Hopenhayn, H. (1992). Entry, exit and firm dynamics in long-run equilibrium. *Econometrica 60*, 1127–1150.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5-6), 953–969.

İmrohoroğlu, A. (1989). Cost of Business Cycles with Indivisibilities and Liquidity Constraints. *Journal of Political Economy 97*(6), 1364–1383.

Inter-American Development Bank (2016). *Saving for Development: How Latin America and the Caribbean Can Save More and Better* (Eduardo Cavallo and Tomás Serebrisky ed.). Development in the Americas. Inter-American Development Bank.

Iregui-Bohórquez, A. M., L. A. Melo-Becerra, M. T. Ramírez-Giraldo, and A. M. Tribín-Uribe (2018). Determinants of Formal and Informal Saving in Colombia. In M. J. Roa and D. Mejía (Eds.), *Financial Decisions of Households and Financial Inclusion: Evidence for Latin America and the Caribbean*, Chapter 4, pp. 95–123. Centro de Estudios Monetarios Latinoamericanos.

Kaboski, J. and R. Townsend (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica 79*(5), 1357–1406.

Kalemli-Ozcan, S., B. Sorensen, and S. Yesiltas (2012). Leverage across firms, banks, and countries. *Journal of International Economics 88*(2), 284–298.

Kaplan, S. N. and L. Zingales (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics 112*(1), 169–215.

Karlan, D. S., A. Lakshmi Ratan, and J. Zinman (2014). Savings by and for the poor: A research review and agenda. *Review of Income and Wealth 60*(1), 36–78.

Karpowicz, I. (2014). Financial Inclusion, Growth and Inequality: A Model Application to Colombia. IMF Working Paper 14/166, International Monetary Fund.

Kast, F. and D. Pomeranz (2014). Saving More to Borrow Less: Experimental Evidence from Access to Formal Savings Accounts in Chile. NBER Working Paper 20239, National Bureau of Economic Research.

Kim, M., D. Kliger, and B. Vale (2003). Estimating switching costs: the case of banking. *Journal of Financial Intermediation 12*(1), 25–56.

Klein, B. (1974). Competitive Interest Payments on Bank Deposits and the Long-Run Demand for Money. *American Economic Review 64*(6), 931–949.

Lagakos, D., M. Mobarak, and M. Waugh (2018). The Welfare Effects of Encouraging Rural-Urban Migration. Working Papers 2018-002, Human Capital and Economic Opportunity Working Group.

Lucas, R. E. (1987). *Models of Business Cycles*. Basil Blackwell.

López, M. R., J. D. Prada, and N. Rodríguez N. (2008). Financial Accelerator Mechanism in a Small Open Economy. Borradores de Economia 525, Banco de la Republica de Colombia.

Mejia, L. F. (2010). Entrepreneurship & interest rate shocks in a small open economy. Technical report, PhD Dissertation, University of Chicago.

Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review 104*(2), 422–458.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review 104*(10), 3186–3221.

Paez, J. A. and C. E. Tamayo (2019). Firm Leverage in Colombia: A New Data Set. Unpublished manuscript, Inter-American Development Bank.

Prada, J. D. and L. E. Rojas (2010). La elasticidad de Frisch y la transmisión de la política monetaria en Colombia. In M. Jalil and L. Mahadeva (Eds.), *Mecanismos de transmisión de la política monetaria en Colombia*, pp. 643–699. Banco de la República - Universidad Externado de Colombia.

Prina, S. (2015). Banking the poor via savings accounts: Evidence from a field experiment. *Journal of Development Economics 115*, 16–31.

Quadrini, V. (2004). Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics 51*(4), 713–751.

Quadrini, V. (2009). Entrepreneurship in macroeconomics. *Annals of Finance 5*(3), 295–311.

Reddy, R., M. Bruhn, and C. Tan (2013). *Financial capability in Colombia: results from a national survey on financial behaviors, attitudes, and knowledge*. The World Bank.

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707–720.

Roa, M. J. and O. Carvallo (2018). *Inclusión financiera y el costo del uso de instrumentos financieros formales: Las experiencias de América Latina y el Caribe*. Banco Interamericano de Desarrollo.

Rojas-Suarez, L. and M. A. Amado (2014). Understanding Latin America's financial inclusion gap. Working Paper 367, Center for Global Development.

Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium models of business cycles. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, pp. 294–330. Princeton University Press.

Solo, T. M. and A. Manroth (2006). Access to financial services in Colombia: the "unbanked" in Bogotá. Policy Research Working Paper 3834, The World Bank.

Wang, F. (2019). An empirical equilibrium model of formal and informal credit markets in developing countries. Unpublished manuscript, University of Houston.

Zuleta, H., J. Parada, A. García, and J. Campo (2010). Participación factorial y contabilidad del crecimiento económico en Colombia. Una propuesta de modificación del método de contabilidad del crecimiento. *Desarrollo y Sociedad 65*, 71–121.

A Appendix

A.1 Firm-level Data

We now briefly summarize the data used for computing the statistics on saving behavior by Colombian firms. A detailed explanation of the dataset is in Paez and Tamayo (2019).

The data is taken from three different sources. First, we collect balance sheet and income data from Colombia's Superintendencia de Sociedades ("Supersociedades") which includes a total of 44,703 privately held firms over the period 2004-2014. These are mostly industrial, medium and large firms which, because of their size or some other industry specific characteristics, are required to file reports to the Superintendency. The dataset includes very detailed account data, including different types of assets (e.g., cash and equivalent, fixed assets, intangible assets, etc.) and liabilites (e.g., short/long-term, with suppliers, banks, etc.), as well as detailed sources of income and expenses.

Next, we use a proprietary database from ORBIS Americas –collected and maintained by Bureau van Dijk– which contains some 474,400 firms over the same period. Unlike the data from Supersociedades, this dataset includes small and often times micro firms from all sectors of the economy. However, we find that for most firms less detail is available in this dataset; in some cases, only total assets or total liabilities (instead of their composition) is available. Finally, the data is complemented by the 47 firms that issue securities and thus report to Colombia's Superintendencia Financiera.

The change in net worth is defined as the yearly change in the difference between total assets and total liabilities, while the firm's "income" is taken to be operating revenue. The former is deflated using the investment deflator from the national accounts, while the latter is deflated at the 2-digit industry level with base year 2015. For the period 2010-2013 –used for the statistics presented in Table 3, we end up with 232,222 firms. Finally, for the calculations used in the calibration, we pick the year 2013 and end up with a sample of some 103,000 firms.

A.2 Exogenous Growth and De-trending

We now present the details of the model under exogenous growth. Recall that an aggregate efficiency component, A_t, is included in the production function, and that it grows deterministically at a constant rate g, $A_t = gA_{t-1}$. This implies that most aggregates in this economy are non-stationary with a deterministic trend. Normalizing $A_0 = 1$ and defining $\gamma = g^{1/(1-\alpha)}$ with $\alpha = \lambda\theta$, such trend becomes γ^t. Denote x as the de-trended value of $\tilde{X}$.

Then the programming problems before de-trending can be written as follows:
Workers:

$$\max_{\tilde{C}_t,\tilde{S}_{t+1},\tilde{Q}_{t+1}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{\tilde{C}_t^{1-\chi}}{1-\chi}$$

subject to

$$\tilde{C}_t + \tilde{Q}_{t+1} + \tilde{S}_{t+1} + \tilde{\xi}\mathbb{I}_{Q_{t+1}>0} = \tilde{W}_t \exp(\varepsilon_t) + (1+r_t)\tilde{Q}_t + \tilde{S}_t.$$

Entrepreneurs:

$$\max_{\tilde{C}_t,\tilde{K}_t,\tilde{B}_{t+1},l_t} E_0 \sum_{t=0}^{\infty} (\beta\eta)^t \frac{\tilde{C}_t^{1-\chi}}{1-\chi}$$

subject to

$$\tilde{C}_t + \tilde{B}_{t+1} = A_t \exp(z_t)^{1-\theta} (\tilde{K}_t^{\lambda} l_t^{1-\lambda})^{\theta} - \tilde{W}_t l_t - (r+\delta)\tilde{K}_t + (1+r)\tilde{B}_t - \tilde{\xi}$$

$$\tilde{K}_t \leq \frac{\tilde{B}_t}{1-\phi}.$$

To save on notation, first define $\alpha = \lambda\theta$ and $\vartheta = (1-\lambda)\theta$ so that output is given by $\tilde{Y}_t = A_t \exp(z_t)^{1-\theta} \tilde{K}_t^{\alpha} l_t^{\vartheta}$. In a balanced growth path, $\tilde{Y}$, $\tilde{B}$, $\tilde{W}$, $\tilde{S}$, $\tilde{Q}$, $\tilde{C}$, and $\tilde{K}$ exhibit a common trend. To find this common trend, recall that $A_t = A_{t-1}g = A_0 g^t$. Normalizing $A_0 = 1$, the common trend can be found to be $g^{(1/(1-\alpha))t}$. To see this, divide by this factor:

$$\frac{\tilde{Y}_t}{g^{(1/(1-\alpha))t}} = \frac{g^t \exp(z_t)^{1-\theta} \tilde{K}_t^{\alpha} l_t^{\vartheta}}{g^{(1/(1-\alpha))t}}.$$

Now, re-write $g^{(1/(1-\alpha))t}$ as $g^{(\alpha/(1-\alpha))t} g^t$ so that

$$\frac{\tilde{Y}_t}{g^{(1/(1-\alpha))t}} = \exp(z_t)^{1-\theta} \left(\frac{\tilde{K}_t}{g^{(1/(1-\alpha))t}} \right)^{\alpha} l_t^{\vartheta}.$$

Hence $\tilde{Y}_t = y_t g^{(1/(1-\alpha))t}$ and $\tilde{K}_t = k_t g^{(1/(1-\alpha))t}$. To save on notation, we define $\gamma = g^{(1/(1-\alpha))}$ so that the common trend is γ^t, but recall that g (not γ) is the rate at which TFP grows. When all trending variables are divided by γ^t, the problems for workers and entrepreneurs can be stated in terms de-trended variables. More specifically, the worker's problem can be written recursively as:

$$W(q,s,\varepsilon) = \max_{c,q',s'} \frac{c^{1-\chi}}{1-\chi} + \beta\gamma^{1-\chi} \sum_{\varepsilon'} W(q',s',\varepsilon')\psi(\varepsilon'|\varepsilon)$$

subject to

$$c + \gamma q' + \gamma s' + \xi \mathbb{I}_{q'>0} = w \exp(\varepsilon) + (1+r)q + s,$$

while the entrepreneur's problem may be written:

$$V(b,z) = \max_{b',k,l} \frac{c^{1-\chi}}{1-\chi} + \beta\eta\gamma^{1-\chi} \sum_{z'} V(b',z')\pi(z'|z)$$

subject to

$$c + \gamma b' + \xi = [\exp(z)]^{1-\theta}(k^{\lambda}l^{1-\lambda})^{\theta} - (r+\delta)k - wl + (1+r)b$$

and the collateral constraint

$$d \leq \varphi k,$$

A.3 Additional Policy Experiments

We now present two additional policy experiments with our baseline model. First, we calibrate the model to resemble the U.S. financial system as a whole instead of matching its saving and credit market outcomes separately. That is, rather than altering τ while keeping ϕ constant at the "Colombia" calibrated level and viceversa, we obtain values for (τ, ϕ) that simultaneously match the informal saving fraction and the credit to output ratio. This experiment requires $(\tau, \phi) = (0.046, 0.536)$. The results are presented in the column labeled "τ = US, ϕ= US" of Table 12 and, as should be expected, are better –in terms of welfare gains– than either policy of "τ = US, ϕ= COL" or "τ= COL, ϕ= US" only, but worse than those presented in the last column of Table 7 where $(\tau, \phi) = (0.045, 0.542)$.

A second experiment that we conduct aims at calibrating the model to match the financial outcomes of Sweden, an economy with very developed financial markets: almost 150 percent of credit to firms as a share of GDP, and an approximate equivalent informal saving of 10 percent (again, using the Findex and ELCA data for Colombia to obtain an equivalent figure for Sweden). This experiment requires $(\tau, \phi) = (0.067, 0.77)$. Such calibration shows that moving toward financial market outcomes that resemble Sweden rather than Colombia brings about very large gains in terms of output, productivity and welfare (on average of up to 20 percent of equivalent increases in consumption for workers and 16 percent for entrepreneurs).

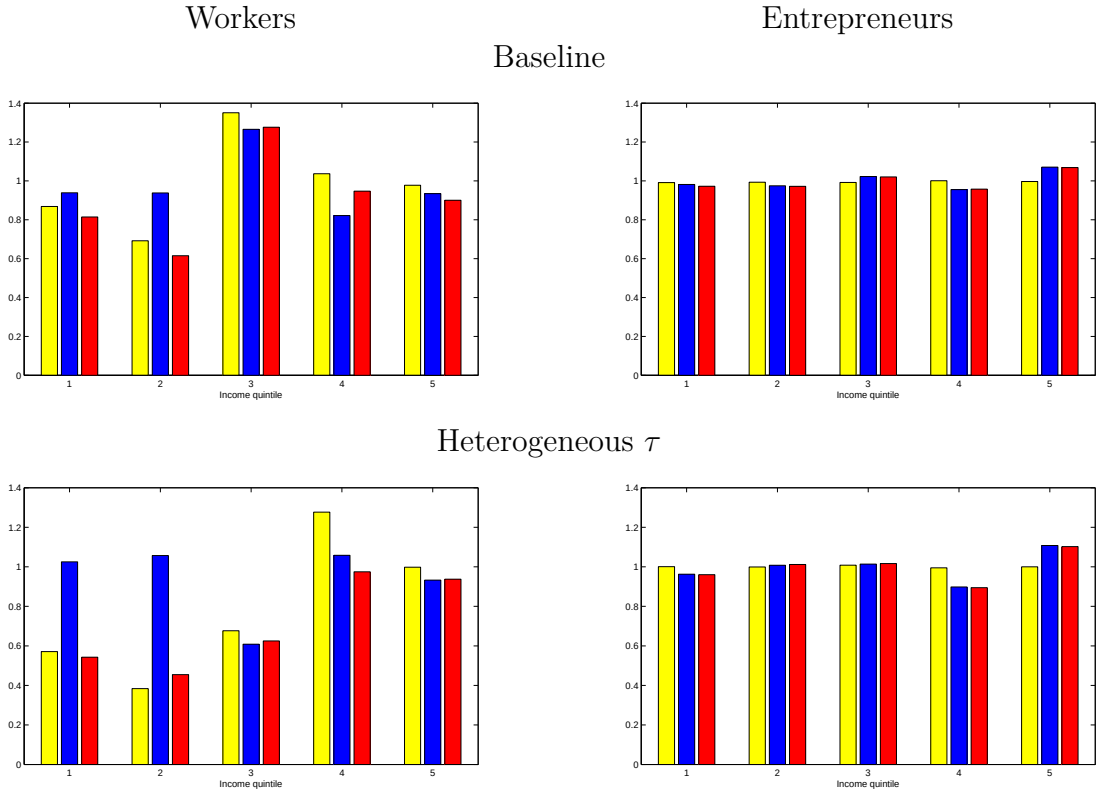
Table 12: Additional policy experiments: US and Sweden

Statistic	Model "Colombia"	τ = US ϕ= US	τ = SWE ϕ= SWE
% of workers who do not save	69.6%	56.8%	49.1%
% of formal savers	62.1%	91.5%	91.0%
Workers' saving rate (savers)	10.8%	16.5%	18.8%
% of entrepreneurs who do not save	31.1%	34.9%	48.3%
Entrepreneurs saving rate (savers)	8.5%	7.2%	7.1%
Credit-to-output ratio	31.6%	95.5%	149.0%
% income in quintiles 1 & 2	8.0%	9.9%	13.0%
% of capital financed by firms	81.8%	51.5%	32.3%
Capital intensity (K/Y)	1.74	1.97	2.2
Aggregate output	1.0	1.10	1.19
Total factor productivity (TFP)	1.0	1.02	1.04
Net real interest rate	8.1%	8.4%	8.4%
Real wage rate	0.37	0.41	0.45
Consumption volatility			
Households (workers)	0.43	0.41	0.40
Firms (entrepreneurs)	0.49	0.46	0.42
Welfare (consumption eq. variation)			
Households (workers)		9.8%	19.7%
Firms (entrepreneurs)		15.5%	15.5%

A.4 Consumption Volatility

As mentioned in the main text, the three policy experiments –reductions in saving constraints, credit frictions, and a combination of policies– result in modest reductions in consumption volatility for low income workers. Figure 5 presents coefficients of variation associated with the policy experiments relative to the benchmark economy.

Figure 5: Consumption Volatility by Income Quintile



Note: The figure shows the distribution of consumption volatility, measured by the coefficient of variation, associated with the policy experiments relative to the benchmark economy. The top panel corresponds to the baseline model, while the bottom panel refers to the model with heterogeneous saving costs.

A.5 An Occupational Choice Model

We now present a simple occupational choice version of the model used in Section 3.

The economy is composed of a measure one of agents who at times may be workers or entrepreneurs. In each period, an agent wakes up to an occupation and makes production and consumption decisions, as well as a choice of occupation for the following period. Agents face idiosyncratic ability shocks captured by x_t, which evolves over time according to a finite-state Markov process with transition probabilities $\mu(x', x) = Pr(x_{t+1}|x_t)$ and ergodic distribution $\Gamma(x)$.

An agent who wakes up to a worker inelastically supplies her unit of labor to an entrepreneur and receives her labor income, $w \exp(x)$, while in the latter case the agent decides how much labor and capital to use in production. Agents also decide how much assets to accumulate. In the case of a worker, she may hold cash, s, or deposits, a, by

paying the fixed cost of using the bank, $\xi = \tau w$. As before, an entrepreneur always pay the fixed cost of using the financial system, but his deposits may be used as collateral to install capital in line with the borrowing constraint of Section 3, $k \leq a(1-\phi)^{-1}$. Entrepreneurs also pay a fixed cost, ϱ, in every period in which the firm operates.

For the case of an agent who wakes up to a worker in the current period, the recursive formulation of the problem can be stated as:

$$W(a,s,x) = \max_{c,a',s'} \frac{c^{1-\chi}}{1-\chi} + \beta \sum_{x'} \mathcal{O}(a',s',x')\mu(x',x)$$

subject to

$$c + \gamma a' + s' + \xi \mathbb{I}_{a'>0} = w \exp(x) + (1+r)a + s,$$

where $\mathcal{O}(a',s',x') = \max\{W(a',s',x'), V(a',s',x')\}$, and $W(a',s',x')$ and $V(a',s',x')$ are the continuation values of being a worker and an entrepreneur, respectively.

For the case of an agent who wakes up to an entrepreneur, the recursive formulation of the problem can be stated as:

$$V(a,s,x) = \max_{b',k,l} \frac{c^{1-\chi}}{1-\chi} + \beta\eta \sum_{x'} \mathcal{O}(a',s',x')\mu(x',x)$$

subject to

$$c + s' + a' + \xi = [\exp(x)]^{1-\theta}(k^{\lambda}l^{1-\lambda})^{\theta} - (r+\delta)k - f_e - wl + (1+r)b + s,$$

and the collateral constraint

$$k \leq a(1-\phi)^{-1}.$$

Notice that, while entrepreneurs are in principle allowed to save in cash, the assumption that they always pay ξ implies that they will never choose to do so. Notice also that an agent that is currently a worker may use her formal asset holdings (a') as collateral for borrowing capital and producing tomorrow (after choosing to become an entrepreneur). The definition of equilibrium would be very similar to that of Section 3, with appropriate changes in notation and with the addition of a policy rule specifying the occupation decision as a function of the states.