Economic Development and the Equilibrium Interaction of Financial Frictions*

Benjamin Moll
Princeton

Robert M. Townsend
MIT

Victor Zhorin
University of Chicago

August 6, 2014

Abstract

Given that micro data imply that the type of financial frictions varies across regions/sectors within a given country, we develop a general equilibrium framework that encompasses these underpinnings. We study the macroeconomic implications of a moral hazard problem due to unobserved effort and contrast them with those of limited commitment, the friction most studied in recent literature. The effects of moral hazard on aggregate productivity and the equilibrium interest rate differ dramatically from those of limited commitment, and the effect of moral hazard and limited commitment within the same economy gives rise to interesting interactions, e.g., interregional capital and labor flows.

*We thank Fernando Aragon, Paco Buera, Hal Cole, Matthias Doepke, Mike Golosov, Cynthia Kinnan, Tommaso Porzio, Yuliy Sannikov, Martin Schneider, Yongs Shin, Ivan Werning and seminar participants at the St Louis Fed, Wisconsin, Northwestern and the Philadelphia Fed for very useful comments. Hoai-Luu Nguyen and Hong Ru provided outstanding research assistance. For sharing their code, we are grateful to Paco Buera and Yongs Shin. Townsend gratefully acknowledges research support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) (grant number R01 HD027638), the research initiative 'Private Enterprise Development in Low-Income Countries' [(PEDL), a programme funded jointly by the Centre for Economic Policy Research (CEPR) and the Department for International Development (DFID), contract reference MRG002_1255], the John Templeton Foundation (grant number 12470), and the Consortium on Financial Systems and Poverty at the University of Chicago (funded by Bill & Melinda Gates Foundation under grant number 51935). The views expressed are not necessarily those of CEPR or DFID. Previous versions of this paper were circulated under the titles "Finance and Development: Limited Commitment vs. Moral Hazard" and "Financial Obstacles and Inter-Regional Flow of Funds."

1 Introduction

There is evidence that even within a given economy, individuals face different types of financial frictions depending on location. In a companion paper, Karaivanov and Townsend (2014) estimate the financial/information regime in place for households including those running businesses in Thailand and find that a moral hazard constrained financial regime fits best in urban areas and a more limited savings regime in rural areas. Similarly, Paulson, Townsend and Karaivanov (2006) using other data argue that moral hazard fits best in the more urban Central region but not in the more rural Northeast. And Ahlin and Townsend (2007) with alternative data find that information seems to be a problem in the Central area, limited commitment in the Northeast.[1]

We begin the next step in this paper and ask what difference the micro financial foundations make for the macro economy and for its macro financial structure. To this end, we develop a general equilibrium framework that encompasses different types of frictions. We then study the macroeconomic implications of a moral hazard problem due to unobserved effort and contrast them with those of limited commitment, the friction most studied in recent literature. Our first major result is that the effects of moral hazard on macroeconomic aggregates differ dramatically from those of limited commitment. We then argue that when these different financial frictions are present in the same economy, they also interact in interesting, unexpected ways, e.g., give rise to interregional capital and labor flows that resemble rural-urban patterns observed in the data. This is our second major result.

In our theory, a large number of households access the economy's capital market via intermediaries with which they form long-term contracts.[2] Households choose between being entrepreneurs and workers. Entrepreneurs produce using labor and capital, and their productivity depends on their talent and a residual productivity term which partly depends on their effort. The intermediary can help finance their capital and potentially insure them against residual productivity risk, but not their talent.[3] Workers supply efficiency units of labor to the economy-wide labor market. Wages depend on their effort and a residual random component which is in principle insurable as with firms. The interest rate and wages are determined in general equilibrium. Contracts between intermediaries and households are subject to one of two frictions: moral hazard or limited commitment. Again, our methods may allow a combination of both. In the moral hazard regime, effort of both entrepreneurs and workers is unobserved so

[1] We discuss these papers, the data used, and the nature of regional variation they find in more detail in Section 5.

[2] Once a household decides to contract with an intermediary, it sticks with that intermediary forever. But the threat of having one's customer poached by another intermediary means that intermediaries have zero net inflows at each point in time.

[3] We exogenously impose the assumption that talent shocks are not insurable in the interest of realism. Our method though would allow us to include these.

that providing full insurance against residual productivity or labor efficiency induces shirking. Households and intermediaries sign optimal dynamic contracts that take this into account. In the limited commitment regime, there is no moral hazard and hence full insurance, but instead entrepreneurs face simple collateral constraints that limit the amount of capital they can use in production to a scalar of their personal wealth.[4]

Our first main result is that the effects of moral hazard on macroeconomic aggregates differ substantially from those of limited commitment.[5] This difference is particularly pronounced for aggregate total factor productivity (TFP) and the equilibrium interest rate. Consider first aggregate TFP. Under moral hazard, TFP is endogenously lower at the firm level because entrepreneurs exert suboptimal effort. This in turns results in depressed aggregate TFP. Moral hazard may also result in a misallocation of capital but this is *not necessarily* the case, a point we emphasize by focussing on an instructive special case in which marginal products of capital are equalized across entrepreneurs.[6] Under limited commitment, in contrast, effort is always chosen optimally and firm-level TFP unaffected, but capital itself is misallocated across firms with given heterogeneous productivities, and this is what depresses aggregate TFP. In a recent paper Midrigan and Xu (2014) have argued that a model with collateral constraints, when calibrated to plant level data, generates fairly small dispersion of marginal products and hence TFP losses from misallocation. The instructive special case of the moral hazard economy just discussed is an example in which there are sizable TFP losses even though returns to capital are equalized across all firms. We next consider the equilibrium interest rate, and find that it is typically lower under limited commitment than under moral hazard. We examine the effects of the two frictions on aggregate capital demand and supply and show that one main reason for this lower interest rate are two well-known results: moral hazard results in individuals being

[4]We choose a formulation of the limited commitment problem that can be represented by a simple static collateral constraint. Alternatively, we could have worked with a more full-blown dynamic limited commitment problem as is common in the optimal contracting literature (for example Albuquerque and Hopenhayn, 2004). We choose to work with collateral constraints, mainly because it facilitates comparison with the existing literature (for example Evans and Jovanovic, 1989; Holtz-Eakin, Joulfaian and Rosen, 1994; Banerjee and Duflo, 2005; Paulson, Townsend and Karaivanov, 2006; Jeong and Townsend, 2007; Buera and Shin, 2013; Moll, 2014; Midrigan and Xu, 2014), and it also simplifies some of the computations.

[5]Of course, the implications of the two frictions also differ at the micro level. But most of these differences are already well understood from the existing literature (see e.g. Rogerson, 1985; Phelan and Townsend, 1991; Albuquerque and Hopenhayn, 2004; Clementi and Hopenhayn, 2006, and the work cited in the first paragraph and section 5.2 below).

[6]In this special case only entrepreneurial effort is unobserved. In contrast capital stocks can be observed and a change in an entrepreneur's capital stock does not change his incentive to shirk. Departures from these assumptions mean that moral hazard leads to a misallocation of capital in addition to suboptimal effort and depressed firm-level TFP. While we think that these are reasonable assumptions – e.g. surely it is easier to observe an entrepreneur's machines than his effort – the main purposes of focussing on this special case is to illustrate in a transparent fashion that moral hazard does not *necessarily* result in capital misallocation.

saving constrained whereas limited commitment induces borrowing constraints.[7]

To obtain our second main result, we study an economy with different frictions within the same economy: a certain exogenous fraction of the population is subject to moral hazard and the remainder is subject to limited commitment. We show that the two frictions interact and that the combined effect of moral hazard and limited commitment is more than just a linear combination of their individual effects. More precisely, aggregate variables such as total factor productivity can be non-monotonic functions of the fraction of the population subject to moral hazard, and we show that this is due to general equilibrium effects.

In these latter mixed regimes, capital and labor flow from the limited commitment to the moral hazard sector, and both income and external finance are higher in the latter. This is due to the difference in financial regimes only – everything else is symmetric. Embedding the moral hazard and limited commitment sectors in the same economy also has non-trivial welfare implications, particularly for the moral hazard sector. As already mentioned, Paulson, Townsend and Karaivanov (2006), Ahlin and Townsend (2007) and Karaivanov and Townsend (2014) suggest that moral hazard may be more prevalent in urban areas and limited commitment in rural areas. When we take this finding at face value and interpret the moral hazard regime as urban areas or industrialized regions and the limited commitment regime as rural areas or regions of the country which are less developed, these patterns resemble rural-urban, inter-regional patterns observed in the data.[8] Of course, in practice there are many other factors that distinguish cities from villages and industrialized from agricultural areas so these findings should be interpreted with great caution. We are not arguing there is one and only one factor but rather that heterogeneous regimes can play an important role.

The bottom line is that the behavior of macro aggregates depends on micro financial underpinnings, and that these may interact in unexpected ways. This has important implications for the literature studying the role of financial market imperfections in economic development. Most of the existing literature works with collateral constraints that are either explicitly or implicitly motivated as arising from a limited commitment problem (Evans and Jovanovic, 1989;

[7]That moral hazard induces saving-constrained behavior is an implication of the inverse Euler equation (Rogerson, 1985). As explained in more detail in Section 4, there are also additional and potentially offsetting effects of the two frictions on capital demand and supply. That being said, a lower interest rate under limited commitment is a robust numerical result that is present under a battery of parameterizations we have tried.

[8]Seminal contributions in development economics emphasized rural to urban labor migration, e.g., Lewis (1954) and Harris and Todaro (1970). Within-country capital flows are somewhat harder to document. Using data from Mexico, an ongoing study by the Comisión Nacional Bancaria y de Valores (CNBV, `http://bit.ly/17NId1F`) funded by CFSP in its efforts to improve flow of funds accounts, finds that municipalities (counties) with cities of more than 300,000 inhabitants tend to borrow from municipalities with smaller or no cities. This is consistent with the capital flows that arise in our model. One of the most well known, if poorly documented, cases of rural to urban flows is Indonesia post big bang reforms in 1983. See Maurer and Seibel (2001), but we do not have within-country flow of funds.

Holtz-Eakin, Joulfaian and Rosen, 1994; Banerjee and Duflo, 2005; Jeong and Townsend, 2007; Buera and Shin, 2013; Buera, Kaboski and Shin, 2011; Moll, 2014; Caselli and Gennaioli, 2013; Midrigan and Xu, 2014). In contrast, there are much fewer studies that model financial frictions as arising from moral hazard. Notable exceptions are the early contributions by Aghion and Bolton (1997) and Piketty (1997), and also Ghatak, Morelli and Sjostrom (2001).[9] Related, some papers study environments with asymmetric information and costly state verification (as in Townsend, 1979), but there are again few of these (Castro, Clementi and Macdonald, 2009; Greenwood, Sanchez and Wang, 2010*a*,*b*; Cole, Greenwood and Sanchez, 2012). Finally, Martin and Taddei (2012) study the implications of adverse selection on macroeconomic aggregates, and contrast them with those of limited commitment. But few authors use micro data to discipline their macro models.[10] Even fewer (perhaps none?) use micro data to choose between the myriad of alternative forms of introducing a financial friction into their model. This is a serious shortcoming, and the goal of this paper is to make some progress by studying the macroeconomic implications of different micro financial underpinnings suggested by the micro data.

Again, the microeconomic literature is somewhat more advanced in terms of taking seriously different micro financial underpinnings and trying to distinguish between them in the data. For example, Albuquerque and Hopenhayn (2004) and Clementi and Hopenhayn (2006) argue that moral hazard and limited commitment have different implications for firm dynamics (see also Schmid (2012)). Krueger and Perri (2011) compare and contrast the permanent income hypothesis versus a model of self-insurance with borrowing constraints and conclude the former explains the dynamics of their data better, and Broer (2013) compares a model with self-insurance to one with limited enforcement. Abraham and Pavoni (2005), Doepke and Townsend (2006) and Attanasio and Pavoni (2011) discuss how consumption allocations differ under moral hazard with and without hidden savings versus full information. The paper by Karaivanov and Townsend (2014) we have already discussed makes a related point but focuses on household-firms and distinguishes between the different regimes. Similarly, Kinnan (2012) uses a different metric based on the first order conditions characterizing optimal insurance under moral hazard,

[9] But note that these authors study overlapping generations models whereas we study an economy with long-term contracts between infinitely-lived households and intermediaries. Another difference between our setup and that of Ghatak, Morelli and Sjostrom (2001) is that in their setup, the principal-agent relationship subject to moral hazard is between entrepreneurs and workers whereas in our setup the principals are intermediaries and the agents are both entrepreneurs and workers. The only other paper we are aware of that features inifinitely-lived entrepreneurs and optimal contracts under moral hazard in general equilibrium is by Shourideh (2012). But his framework differs in that moral hazard stems from the unobservability of capital as opposed to effort, there is no occupational choice, and his focus is on optimal taxation of entrepreneurial income as opposed to understanding cross-country income differences. Of course, moral hazard plays a lead role in macro financial literature on regulation. See Kareken and Wallace (1978) onward to the present day.

[10] One exception is Midrigan and Xu (2014).

limited commitment and hidden income to distinguish these regimes in Thai data. Meisenzahl (2011) is another example trying to distinguish between different regimes using micro data, in his case between moral hazard and costly state verification and using data on small businesses in the US.

The paper is organized as follows. Section 2 develops our theory, and section 3 discusses our choice of parameter values. Section 4 compares the macroeconomic effects of moral hazard to those of limited commitment. Section 5 presents results for mixed regimes in which part of the economy is subject to moral hazard and the remainder to limited commitment. Section 6 discusses robustness of our results to alternative parameterizations, and Section 7 is a conclusion.

2 Households and Intermediaries

We consider an economy populated by a continuum of households of measure one, indexed by $i \in [0, 1]$ and a continuum of intermediaries, indexed by j. Time is discrete. In each period t, a household experiences two shocks: an ability shock, z_{it} and an additional "residual productivity" shock, ε_{it} (more on this below). Households have preferences over consumption, c_{it} and effort, e_{it}

$$v_{i0} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_{it}, e_{it}).$$

Households can access the capital market of the economy only via one of the intermediaries. Each intermediary contracts with a continuum of households and therefore also provides some insurance to households. Intermediaries compete ex-ante for the right to contract with households. Once a household i decides to contract with an intermediary j, he sticks with that intermediary forever. At the same time, we assume that intermediaries can poach customers from each other based on their observable characteristics (talent and wealth). This means that for each such group of customers, net resource flows into the intermediary must be zero.

Households have some initial wealth a_{i0} and an income stream $\{y_{it}\}_{t=0}^{\infty}$ (determined below). When households contract with an intermediary, they give their entire initial wealth and income stream to the intermediary. The intermediary pools the income of all the households it contracts with, invests it at a risk-free interest rate r_t, and transfers some consumption to the households. An intermediary together with the continuum of households it contracts with therefore forms a mutual fund or a "risk-sharing group": some of each household's risk is shared with the other households in the group according to an optimal contract specified below. Denote by a_{jt} and y_{jt} the pooled wealth and income in risk-sharing group j (that is, run by intermediary j). Then the risk-sharing group's budget constraint is

$$a_{jt+1} = y_{jt} - c_{jt} + (1 + r_t)a_{jt}. \tag{1}$$

The optimal contract between intermediary and households maximizes the households' utility subject to this budget constraint (and incentive constraints specified below). Because net resource flows into the intermediary must be zero for each group of individuals with the same observed characteristics (here wealth a_{it} and talent z_{it}), this problem is equivalent to maximizing expected utility for each of these groups. Risk-sharing groups make their decisions taking as given current and future time profiles of wages w_t and interest rates r_t respectively and compete with each other in competitive labor and capital markets. Mostly, however, one treats these factor prices as a constant (over time), namely wage and interest rate w and r respectively. We here assume that the economy is in a stationary equilibrium so that factor prices are constant over time. Again, this is mainly for simplicity. Our setup can easily be extended to the case where aggregates vary deterministically over time at the expense of some extra notation.

2.1 Household's Problem

Households can either be entrepreneurs or workers. We denote by $x_{it} = 1$ the choice of being an entrepreneur and by $x_{it} = 0$ that of being a worker. First, consider entrepreneurs. An entrepreneur hires labor ℓ_{it} at a wage w_t and rents capital k_{it} at a rental rate $r_t + \delta$ and produces some output.[11] His observed productivity has two components: a component, z_{it}, that is known by the entrepreneur in advance at the time he decides how much capital and labor to hire, and a residual component, ε_{it}, that is realized afterwards. We will call the first component *entrepreneurial ability* and the second *residual productivity*. The evolution of entrepreneurial talent is exogenous and given by some stationary transition process $\mu(z_{it+1}|z_{it})$. Residual productivity instead depends on an entrepreneur's effort, e_{it}, which is potentially unobserved, depending on the financial regime. More precisely, his effort determines the distribution $p(\varepsilon_{it}|e_{it})$ from which residual productivity is drawn, with higher effort making good realizations more likely. We assume that intermediaries can insure residual productivity ε_{it}. In contrast, even if entrepreneurial ability, z_{it}, is observed, it is *not* contractible and hence cannot be insured. An entrepreneur's output is given by

$$z_{it}\varepsilon_{it}f(k_{it}, \ell_{it}),$$

where $f(k, \ell)$ is a span-of-control production function.

Next, consider workers. A worker sells efficiency units of labor ε_{it} in the labor market at wage w_t. Efficiency units are observed but are stochastic and depend on the worker's

[11] We assume that capital is owned and accumulated by a capital producing sector. This sector rents out capital to entrepreneurs in a capital rental market, and also holds the net debt of households (or more precisely, of the risk-sharing groups the households belong to) between periods. See Appendix B for details. That the rental rate equals $r_t + \delta$ follows from a standard arbitrage argument. This way of stating the problem avoids carrying capital, k_{it}, as a state variable in the dynamic program of a risk-sharing group.

true underlying effort, with distribution $p(\varepsilon_{it}|e_{it})$.[12] The worker's true underlying effort is potentially unobserved, depending on the financial regime. A worker's ability is fixed over time and identical across workers, normalized to unity.

Putting everything together, the income stream of a household is

$$y_{it} = x_{it}[z_{it}\varepsilon_{it}f(k_{it}, \ell_{it}) - w_t\ell_{it} - (r_t + \delta)k_{it}] + (1 - x_{it})w_t\varepsilon_{it}. \tag{2}$$

The joint budget constraint of the risk-sharing group consisting of households and intermediary is given by (1) where y_{jt} is the sum over y_{it} of all households that contract with intermediary j.

The timing is illustrated in Figure 1 and is as follows: the household comes into the period

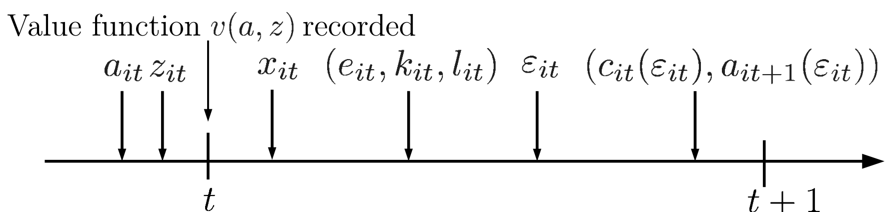


Figure 1: Timing

with previously determined savings a_{it} and a draw of entrepreneurial talent z_{it}. Then within period t, the contract between household and intermediary assigns occupational choice x_{it}, effort, e_{it}, and – if the chosen occupation is entrepreneurship – capital and labor hired, k_{it} and ℓ_{it}, respectively. All these choices are conditional on talent z_{it} and assets carried over from the last period, a_{it}. Next, residual productivity, ε_{it}, is realized which depends on effort through the conditional distribution $p(\varepsilon_{it}|e_{it})$. Finally, the contract assigns the household's consumption and savings, that is functions $c_{it}(\varepsilon_{it})$ and $a_{it+1}(\varepsilon_{it})$. The household's effort choice e_{it} may be unobserved depending on the regime we study. All other actions of the household are observed. For instance, there are no hidden savings.

We now write the problem of a risk-sharing group, consisting of a household and an intermediary, in recursive form. The two state variables are wealth, a, and entrepreneurial ability, z. Recall that z evolves according to some exogenous Markov process $\mu(z'|z)$. It will be convenient below to define the household's expected continuation value by

$$\mathbb{E}_{z'}v(a', z') = \sum_{z'} v(a', z')\mu(z'|z),$$

[12] The assumption that the distribution of workers' efficiency units $p(\cdot|e_{it})$ is the same as that of entrepreneurs' residual productivity is made solely for simplicity, and we could easily allow workers and entrepreneurs to draw from different distributions at the expense of some extra notation.

where the expectation is over z'. A contract between a household of type (a, z) and an intermediary solves

$$v(a,z) = \max_{x,e,k,\ell,c(\varepsilon),a'(\varepsilon)} \sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e] + \beta\mathbb{E}_{z'}v[a'(\varepsilon),z']\right\} \qquad \text{s.t.} \tag{3}$$

$$\sum_{\varepsilon} p(\varepsilon|e)\left\{c(\varepsilon)+a'(\varepsilon)\right\} = \sum_{\varepsilon} p(\varepsilon|e)\left\{x[z\varepsilon f(k,\ell) - w\ell - (r+\delta)k] + (1-x)w\varepsilon\right\} + (1+r)a \tag{4}$$

and also subject to regime-specific constraints specified below.

The contract maximizes a household's expected utility subject to a break-even constraint for the intermediary. This is because competition by intermediaries for households ensures that any intermediary has zero net capital inflows in expectation. Note that the budget constraint of a risk syndicate (4) averages over realizations of ε; it does not have to hold separately for every realization of ε. This is because the contract between the household and the intermediary has an insurance aspect and there are a continuum of households, hence no group aggregate risk. This insurance also implies that consumption at the individual level can be different from income less than savings. Such an insurance arrangement can be "decentralized" in various ways. The intermediary could simply make state-contingent transfers to the household. Alternatively, intermediaries can be interpreted as banks that offer savings accounts with state-contingent interest payments to households.

In contrast to residual productivity ε, talent z is assumed to not be insurable. *Prior* to the realization of ε, the contract specifies consumption and savings that are *contingent* on ε, $c(\varepsilon)$ and $a'(\varepsilon)$. In contrast, consumption and savings cannot be contingent on next period's talent realization z'.[13]

The contract between intermediaries and households is subject to one of two frictions: private information in the form of moral hazard or limited commitment. Each friction corresponds to a regime-specific constraint that is added to the dynamic program (3) and (4). In line with the findings of Paulson, Townsend and Karaivanov (2006), Ahlin and Townsend (2007) and Karaivanov and Townsend (2014) discussed in the introduction, a possible interpretation is that different financial regimes represent different locations or sectors within an economy: moral hazard in urban and industrialized areas and limited commitment in rural and agricultural areas. We will pursue this interpretation in more detail in section 5. For sake of simplicity and to isolate the economic mechanisms at work, the only thing that varies across the two regimes is the financial friction. It would be easy to incorporate some differences, say in the stochastic processes for ability z and residual productivity ε at the expense of some extra

[13]The above dynamic program could be modified to allow for talent to be insured as follows: allow agents to trade in assets whose payoff is contingent on the realization of next period's talent z'. On the left-hand side of the budget constraint (4), instead of $a'(\varepsilon)$, we would write $a'(\varepsilon, z')$ and sum these over future states z' using the probabilities $\mu(z'|z)$ so that z' does not appear as a state variable next period, as its realization is completely insured and that insurance is embedded in the resource constraint.

notation. We specify the two financial regimes in turn.

2.2 Moral Hazard

In this regime, effort e is unobserved. Since the distribution of residual productivity, $p(\varepsilon|e)$ depends on effort, this gives rise to a standard moral hazard problem: full insurance against residual productivity shocks would induce the household to shirk, to exert suboptimal effort. The contract takes this into account in terms of an incentive-compatibility constraint:

$$\sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e]+\beta\mathbb{E}_{z'}v[a'(\varepsilon),z']\right\} \geq \sum_{\varepsilon} p(\varepsilon|\hat{e})\left\{u[c(\varepsilon),\hat{e}]+\beta\mathbb{E}_{z'}v[a'(\varepsilon),z']\right\} \;\forall e,\hat{e}. \tag{5}$$

This constraint ensures that the value to the household of choosing the effort level assigned by the contract, e, is at least as large as that of any other effort, $\hat{e}$. The optimal dynamic contract in the presence of moral hazard solves (3) subject (4) and the additional constraint (5). As already mentioned, to fix ideas, we would like to think of this regime as representing the prevalent form of financial contracts in urban and industrialized areas.

Relative to existing theories of firm dynamics with moral hazard, our formulation in (5) is special in that only entrepreneurial effort is unobserved. In contrast capital stocks can be observed and a change in an entrepreneur's capital stock does not change his incentive to shirk. More precisely, the distribution of relative output obtained from two different effort levels does not depend on the level of capital. This is a result of two assumptions: that output depends on residual productivity ε in a multiplicative fashion, and that the distribution of residual productivity $p(\varepsilon|e)$ does not depend on capital (i.e. it is not given by $p(\varepsilon|e,k)$). We focus on this instructive special case because – as we will show below – it illustrates in a transparent fashion that moral hazard does not *necessarily* result in capital misallocation but that it can nevertheless have negative effects on aggregate productivity, GDP and welfare.

The literature on optimal dynamic contracts under private information typically makes use of an alternative formulation which uses promised utility as a state variable (Spear and Srivastava, 1987) and features a "promise-keeping" constraint, neither of which are present here. The connection between this formulation and ours is as follows. Consider first a special case with no ability (z) shocks, and only residual productivity (ε) shocks. In this case, the two formulations are equivalent, a result that we establish in Appendix C. In this sense, the insurance arrangement regarding ε-shocks is optimal (again taking all paths of interest rates and wages as fixed) . The equivalence between the two formulations no longer holds in the case with both z-shocks and ε-shocks. This is because we rule out insurance against z-shocks by assumption, whereas an optimal dynamic contract would allow for such insurance.[14] We would

[14] To see the lack of insurance against z-shocks, consider the case where residual productivity shocks are shut down, $\varepsilon = 1$ with probability one. Then our formulation is an income fluctuations problem, like Schechtman and

like to reiterate, however, that we do not limit insurance arrangements regarding ε-shocks, as shown by the equivalence with an optimal dynamic contract in the absence of z-shocks.

When solving the problem (3) to (5) numerically, we allow for lotteries in the optimal contract to "convexify" the constraint set as in Phelan and Townsend (1991). See Appendix D for the statement of the problem (3) to (5) with lotteries.

2.3 Limited Commitment

In this regime, effort e is observed. Therefore, there is no moral hazard problem and the contract consequently provides perfect insurance against residual productivity shocks, ε. Instead we assume that the friction takes the form of a simple collateral constraint:

$$k \leq \lambda a, \quad \lambda \geq 1. \tag{6}$$

This form of constraint has been frequently used in the literature on financial frictions (see, for example, Evans and Jovanovic, 1989; Holtz-Eakin, Joulfaian and Rosen, 1994; Banerjee and Duflo, 2005; Paulson, Townsend and Karaivanov, 2006; Buera and Shin, 2013; Moll, 2014; Midrigan and Xu, 2014). It can be motivated as a limited commitment constraint.[15] The exact form of the constraint is chosen for simplicity. Some readers may find it more natural if the constraint were to depend on talent $k \leq \lambda(z)a$ as well. This would be relatively easy to incorporate, but others have shown that this affects results mainly quantitatively but not qualitatively (Buera, Kaboski and Shin, 2011; Moll, 2014). The assumption that talent z is stochastic but cannot be insured makes sure that collateral constraints bind for some individuals at all points in time. If instead talent were fixed over time for example, individuals would save themselves out of collateral constraints over time (Banerjee and Moll, 2010).

The optimal contract in the presence of limited commitment solves (3) subject to (4) and the additional constraint (6). As already mentioned, to fix ideas and in line with Paulson, Townsend and Karaivanov (2006), Ahlin and Townsend (2007) and Karaivanov and Townsend

Escudero (1977), Aiyagari (1994) or other Bewley models. One reason we rule out insurance against z-shocks is that this assumption allows for a determinate stationary wealth distribution in the absence of moral hazard or limited commitment. In that case, if z-shocks were insurable, the economy would aggregate to a neoclassical growth model and in steady state only aggregate wealth (but not its distribution) would be determined. That being said, in principle, we could handle insurance against z shocks as described in footnote 13.

[15] Consider an entrepreneur with wealth a who rents k units of capital. The entrepreneur can steal a fraction $1/\lambda$ of rented capital. As a punishment, he would lose his wealth. In equilibrium, the financial intermediary will rent capital up to the point where individuals would just be on the margin of having an incentive to steal the rented capital, implying a collateral constraint $k/\lambda \leq a$ or $k \leq \lambda a$. Alternatively, we could have worked with a more full-blown dynamic limited commitment problem as is common in the optimal contracting literature (for example Albuquerque and Hopenhayn, 2004). We choose to work with collateral constraints, mainly because it facilitates comparison with the existing literature, and it also simplifies some of the computations.

(2014), we would like to think of this regime as capturing the workings of financial markets in rural and agricultural areas.

2.4 Factor Demands and Supplies

Risk-sharing groups interact in competitive labor and capital markets, taking as given the sequences of wages and interest rates. Denote by $k(a, z; w, r)$ and $\ell(a, z; w, r)$ the common (across risk-sharing groups) optimal capital and labor demands of households with current state (a, z). A worker supplies ε efficiency units of labor to the labor market, so labor supply of a cohort (a, z) is

$$n(a, z; w, r) \equiv [1 - x(a, z)] \sum_{\varepsilon} p(\varepsilon|e(a, z))\varepsilon. \tag{7}$$

Note that we multiply by the indicator for being a worker, $1 - x$, so as to only pick up the efficiency units of labor by the fraction of the cohort who decide to be workers. Finally, individual capital supply is simply a household's wealth, a.

2.5 Equilibrium

We use the saving policy functions $a'(\varepsilon)$ and the transition probabilities $\mu(z'|z)$ to construct transition probabilities $\Pr(a', z'|a, z)$. In the computations we discretize the state space for wealth, a, and talent, z, so this is a simple Markov transition matrix. Given these transition probabilities and an initial distribution $g_0(a, z)$, we then obtain the sequence $\{g_t(a, z)\}_{t=0}^{\infty}$ from

$$g_{t+1}(a', z') = \Pr(a', z'|a, z)g_t(a, z). \tag{8}$$

Note that we cannot guarantee that the process for wealth and ability (8) has a unique and stable stationary distribution. While the process is stationary in the z-dimension (recall that the process for z, $\mu(z'|z)$, is exogenous and a simple stationary Markov chain), the process may be non-stationary or degenerate in the a-dimension. That is, there is the possibility that the wealth distribution either fans out forever or collapses to a point mass. Similarly, there may be multiple stationary equilibria. In the examples we have computed, these issues do however not seem to be a problem and (8) always converges, and from different initial distributions.

Once we have found a stationary distribution of states from (8), we check that markets clear and otherwise iterate. Denote the stationary distribution of ability and wealth by $G(a, z)$. Then market clearing is

$$\int \ell(a, z; w, r)dG(a, z) = \int n(a, z; w, r)dG(a, z), \tag{9}$$

$$\int k(a, z; w, r)dG(a, z) = \int a dG(a, z). \tag{10}$$

The equilibrium factor prices w and r are found using the algorithm outlined in Appendix A.1 of Buera and Shin (2013).

3 Parameterization

The next section presents some numerical results. The present section discusses the functional forms and parameter values we use when computing these.

Functional forms We assume that utility is separable and isoelastic

$$u(c,e) = U(c) - V(e), \quad U(c) = \frac{c^{1-\sigma}}{1-\sigma}, \quad V(e) = \frac{\chi}{1+1/\varphi} e^{1+1/\varphi}, \tag{11}$$

and that effort, e, can take values in some bounded interval $[\underline{e}, \bar{e}]$. The parameter σ is the inverse of the intertemporal elasticity of substitution and also the coefficient of relative risk aversion. The parameter φ is the Frisch elasticity of labor supply. The production function is Cobb-Douglas

$$\varepsilon z f(k,\ell) = \varepsilon z k^{\alpha} \ell^{\gamma}. \tag{12}$$

We assume that $\alpha + \gamma < 1$ so that entrepreneurs have a limited span of control and positive profits. We assume the following transition process $\mu(z'|z)$ for entrepreneurial ability following Buera, Kaboski and Shin (2011) and Buera and Shin (2013): with probability ρ a household keeps his current ability z; with probability $1-\rho$ she draws a new entrepreneurial ability from a discretized version of a truncated Pareto distribution whose CDF is[16]

$$\Psi(z) = \frac{1-(z/\underline{z})^{-\zeta}}{1-(\bar{z}/\underline{z})^{-\zeta}},$$

where $\underline{z}$ and $\bar{z}$ are the lower and upper bounds on ability. We further assume that residual productivity takes two possible values $\varepsilon \in \{\varepsilon^L, \varepsilon^H\}$ and that the probability of the good draw depends on effort as follows:

$$p(\varepsilon^H|e) = (1-\theta)\frac{1}{2} + \theta \frac{e-\underline{e}}{\bar{e}-\underline{e}}.$$

The parameter $\theta \in (0,1)$ controls the sensitivity of the residual productivity distribution with respect to effort (and recall that $\underline{e}$ and $\bar{e}$ are the lower and upper bounds on effort). Note that under full insurance against ε, what matters for the incentive of an agent to exert effort is only θ *relative* to the disutility parameter χ. That is, since χ scales the marginal cost of effort, and θ scales the marginal benefit, what matters is the ratio χ/θ.

[16] The probability distribution of z' conditional on z is therefore $\mu(z'|z) = \rho\delta(z'-z) + (1-\rho)\psi(z')$ where $\delta(\cdot - z)$ is the Dirac delta function centered at z and $\psi(z) = \Psi'(z)$ is the PDF corresponding to Ψ.

Parameter values Table 1 presents the parameter values we use in our numerical experiments. The preference parameters β, σ, φ are set to standard values in the literature.[17]

Table 1: Parameter Values in Benchmark Economy

Parameter	Value	Description
β	1.05^{-1}	discount factor
σ	2	inverse of intertemporal elasticity of substitution
φ	2	Frisch elasticity
χ	0.525	disutility of labor
α	0.3	exponent on capital in production function
γ	0.4	exponent on labor in production function
δ	0.06	depreciation rate
ρ	0.75	persistence of entrepreneurial talent
ζ	1	tail param. of talent distribution (truncated Pareto)
$\underline{z}$	1	lower bound on entrepreneurial talent
$\bar{z}$	4	upper bound on entrepreneurial talent
θ	0.2	sensitivity of residual productivity to effort
ε^L	0	value of low residual productivity draw
ε^H	2	value of high residual productivity draw
λ	1.8	tightness of collateral constraints

As already noted, under full insurance against ε only the ratio $\tilde{\chi} = \chi/\theta$ matters. We set $\tilde{\chi} = \chi/\theta = 2.625$, which lies in the range usually considered in the literature.[18] We set the sensitivity of residual productivity ε with respect to effort e equal to $\theta = 0.2$ implying that $\chi = \tilde{\chi}\theta = 2.625 \times 0.2 = 0.525$. We conduct robustness exercises with respect to θ below. We set returns to scale equal to $\alpha + \gamma = 0.7$ which is close to values considered in the literature.[19] The one-year depreciation rate is set at $\delta = 0.06$.

Next consider the parameters governing the ability and residual productivity processes. We set the persistence of entrepreneurial talent to $\rho = 0.75$. This is consistent with empirical estimates (Gourio, 2008; Collard-Wexler, Asker and DeLoecker, 2011), and similar to the parameter value used by Midrigan and Xu (2014) (0.74, see their Table 2). We set the tail parameter of

[17]Perhaps the most challenging among these is the Frisch elasticity φ. For instance Shimer (2010) argues that a range of 1/2 to 4 covers most values that either micro- and macroeconomists would consider reasonable ($\varphi = 4$ corresponds to the value in Prescott (2004)). Our choice of $\varphi = 2$ is in the middle of this range.

[18]The macroeconomics literature typically assumes that $\theta = 1$ so that effort translates one for one into efficiency units of labor. In this case $\tilde{\chi} = \chi$ and only this utility shifter has to be calibrated. See for example Prescott (2004) and Shimer (2010) who use a similar value for $\tilde{\chi}$ as we do.

[19]For example, Buera, Kaboski and Shin (2011) and Buera and Shin (2013) set returns to scale equal to 0.79.

the talent distribution to $\zeta = 1$ which would correspond to Zipf's law if the Pareto distribution were unbounded. We normalize the lower bound of talent to $\underline{z} = 1$, and set the upper bound four times higher, $\bar{z} = 4$. This talent range is in line with that typically considered in the literature (for example Buera and Shin, 2013; Buera, Kaboski and Shin, 2011, although their Pareto distributions feature thinner tails). In any case, we argue in section 6 below that our results are robust to a variety of alternative parameterizations. Finally, for our benchmark numerical results, we set the parameter λ governing the tightness of the collateral constraints, equation (6), to $\lambda = 1.8$. In our limited commitment economy, this results in an external finance to GDP ratio of 1.468 which is close to the values of the 2007 external finance to GDP ratios of Brazil (1.366), China (1.463), India (1.588) and Thailand (1.676).[20]

4 Moral Hazard vs. Limited Commitment

In this section we study the macroeconomic implications of the moral hazard problem outlined in section 2.2, and argue that these differ substantially from those of the limited commitment problem in section 2.3. Table 2 shows the effects of the two frictions on various macroeconomic aggregates. We here emphasize two of these: aggregate productivity and the equilibrium interest rate.

Table 2: Comparison of Different Regimes

	Moral Hazard	Limited Commitment
TFP (% of FB)	0.885	0.872
GDP (% of FB)	0.859	0.777
Welfare (% of FB)	0.891	0.592
Wage (%of FB)	0.931	0.764
Interest Rate	0.010	-0.041
% Entrepreneurs	0.143	0.227

Consider first aggregate TFP reported in the first row, which is normalized by its value in a frictionless economy without constraints (5) or (6).[21] Both moral hazard (column 1) and limited commitment (column 2) result in TFP losses of roughly twelve percent. However, and as we explain in more detail below, the mechanism through which these TFP losses occur is quite different. Under moral hazard, TFP is endogenously lower at the firm level because entrepreneurs exert suboptimal effort, and this in turn depresses aggregate TFP. In the special case of the moral hazard formulation we analyze, this is the only source of TFP losses. Under

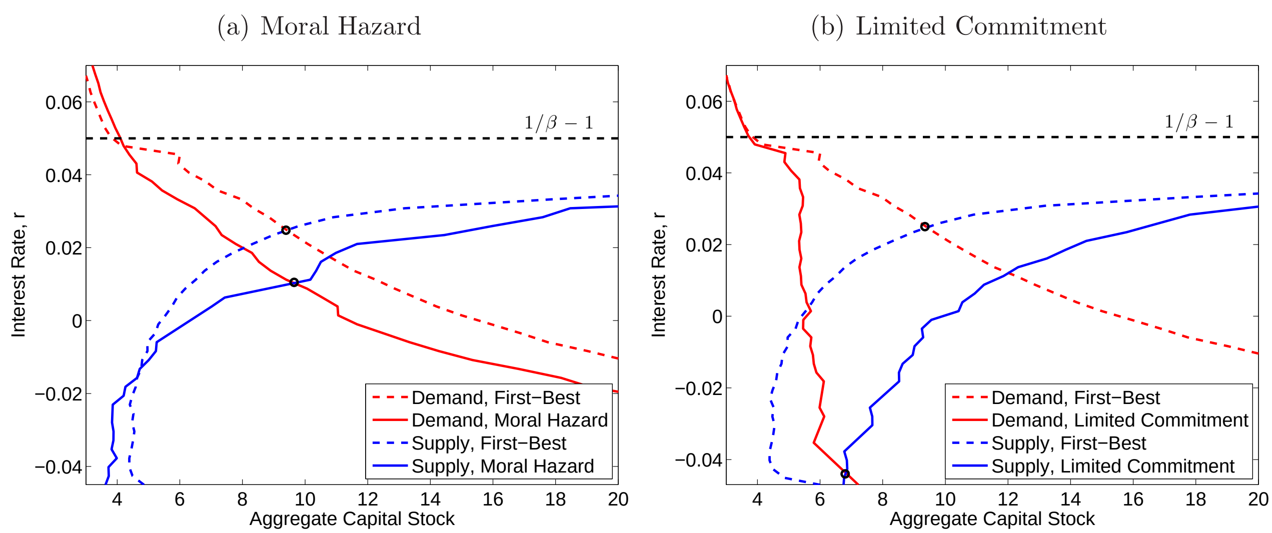
[20]These numbers are from Beck, Demirguc-Kunt and Levine (2000). External finance is defined to be the sum of private credit, private bond market capitalization, and stock market capitalization. This definition follows Buera, Kaboski and Shin (2011). See also their footnote 9.

[21]Aggregate TFP is computed as $TFP = Y/(K^{\nu}L^{1-\nu})$ where Y is aggregate output, K is the aggregate capital stock, L is aggregate labor and $\nu = \frac{\alpha}{\alpha+\gamma}$.

limited commitment, in contrast, effort is always chosen optimally and firm-level TFP unaffected, but capital itself is misallocated across firms with given heterogeneous productivities. This effect of limited commitment on aggregate TFP has been discussed in various previous studies (e.g., Buera, Kaboski and Shin, 2011; Buera and Shin, 2013). We will return to the differential implications of the two frictions for aggregate TFP below, when we examine their effects at the firm-level.

Next, consider the equilibrium interest rate. The interest rate is depressed relative to the rate of time preference in both regimes. But it is considerably lower in the limited commitment economy than in the moral hazard economy.[22] This finding is an important building block for our analysis in Section 5 of the equilibrium interaction of different financial frictions within the same economy. We therefore explore it in more detail in Figure 2 which graphically examines how the aggregate demand for and supply of capital determine the equilibrium interest rate (as in Aiyagari, 1994). Panel (a) plots capital demand and supply for the moral hazard regime

Figure 2: Determination of Equilibrium Interest Rate



(solid lines) and contrasts them with demand and supply in the "first-best" economy without moral hazard (dashed lines). For each value of the interest rate, the wage is recalculated so as to clear the labor market. Panel (b) repeats the same exercise for the limited commitment regime. The first-best demand and supply (the dashed lines) are the same in the two panels and serve as a benchmark to assess the differential effects of the two frictions on the interest rate.

[22]Some readers may wonder about its level, namely why real interest rates are negative. Interest rates are bounded below by $-\delta$ and negative *real* interest rates due to depressed credit demand are a common feature of models with collateral constraints (Buera and Shin, 2013; Buera, Kaboski and Shin, 2011; Guerrieri and Lorenzoni, 2011). That being said, many alternative parameterizations (in particular those with lower discount factor β) feature positive interest rates.

Consider first the moral hazard economy in panel (a). Relative to the first-best, moral hazard depresses capital demand for all relevant values of the interest rate. This is because moral hazard results in entrepreneurs and workers exerting suboptimal effort which depresses the marginal productivity of capital. The effect of moral hazard on capital supply is ambiguous and differs according to the value of the interest rate. It turns out that this ambiguity is the result of a direct effect and a counteracting general equilibrium effecting operating through wages. For a given *fixed wage*, moral hazard always *decreases* capital supply, i.e. capital supply shifts to the left. This is due to a well-known result: the inverse Euler equation of Rogerson (1985) which states that the optimal contract under moral hazard discourages saving whenever the incentive compatibility constraint (5) binds and hence results in individuals being *saving constrained* (see also Ligon, 1998; Golosov, Kocherlakota and Tsyvinski, 2003). Lemma 1 in Appendix E.1 derives the appropriate variant of this result for our framework and discusses the intuition in more detail.[23] But counteracting this negative effect on capital supply is a positive general equilibrium effect: labor demand and hence the wage fall, resulting in more entry into entrepreneurship, higher aggregate profits and higher savings.[24] The overall effect is ambiguous, and in our computations capital supply shifts to the right for some values of the interest rate and to the left for others.

Contrast this with the limited commitment economy in panel (b). Under limited commitment, capital demand shifts to the left whereas capital supply shifts to the right. The drop in capital demand is a direct effect of the constraint (6), and it is considerably larger than the demand drop under moral hazard. That capital supply shifts to the right is due to increased self-financing of entrepreneurs (Buera, Kaboski and Shin, 2011; Buera and Shin, 2013, among others). As a result the interest rate drops considerably relative to the first-best, and more so than under moral hazard. Obviously the size of this drop depends on the parameter λ which governs how binding the limited commitment problem is. But our findings are qualitatively unchanged for many different values of λ.[25]

While we are unable to prove analytically that the equilibrium interest rate is lower under limited commitment than under moral hazard, the finding is present in all our numerical experiments and under a big variety of alternative parameterizations (see Section 6). This is not surprising, given that Figure 2 suggests that there are some strong forces pushing in this direction.[26] Foremost among these is that, under moral hazard, individuals are savings

[23]In line with the inverse Euler equation, the finding that the introduction of moral hazard reduces capital supply for a given wage and interest rate is present in all our numerical examples.

[24]Lower wages also lead workers to save less but this effect is negligible in all our computations.

[25]As discussed in Section 3, the value for λ can be disciplined with data on external finance to GDP ratios. That the interest rate under limited commitment is lower than that under moral hazard is true for all values of λ that are consistent with external finance to GDP ratios for low and middle income countries. Also see the discussion of alternative parameterizations in Section 6.

[26]From the Figure, it can be seen that the opposite result, namely that the interest rate is higher under

constrained which, all else equal, pushes up the interest rate; in contrast, limited commitment results in higher savings due to self-financing which pushes down interest rates. Also going in this direction is that in practice, limited commitment results in a greater drop in capital demand than moral hazard.[27]

Of course, other macroeconomic aggregates besides aggregate productivity and the equilibrium interest rate are also affected, and we briefly discuss these here. Under our parameterization, GDP, welfare and the wage rate are more depressed under limited commitment than under moral hazard.[28] This is due to a stronger negative effect of limited commitment on aggregate capital accumulation (see the equilibrium capital stocks in Figure 2). However, as discussed in section 6, under alternative parameter values these variables may be more depressed in the moral hazard regime.

To explore in more detail the effect of moral hazard on the aggregates in Table 2 and the contrast with limited commitment, we now present some numerical results examining variables at the micro level in the stationary equilibrium of our economy. We here focus on distributions of variables that are most helpful for understanding the behavior of aggregates. A detailed discussion of distributions of a number of additional variables of interest is in Appendix E. Figure 3 plots the distribution of entrepreneurial effort under moral hazard and contrasts it with that under limited commitment. Under moral hazard (panel a), a relatively large fraction of entrepreneurs chooses the lowest possible effort level or an effort level close to that.[29] Under

limited commitment than under moral hazard, is a theoretical possibility. This would require that, first, moral hazard results in a greater drop in capital demand than limited commitment; and second, that the general equilibrium effect discussed above is so strong that it results in a large rightward shift of moral hazard capital supply. Neither scenario occurred under any of our alternative parameterizations and a lower interest rate under limited commitment is therefore a robust numerical result.

[27]As already noted in footnote 25, the demand drop under limited commitment is relatively large for values of the parameter λ consistent with external finance to GDP ratios observed in the data. Similarly, the size of the demand drop under moral hazard is always relatively small, except when residual productivity is extremely responsive to individuals' effort choice (both the support $[\varepsilon_L, \varepsilon_H]$ is large and θ is high).

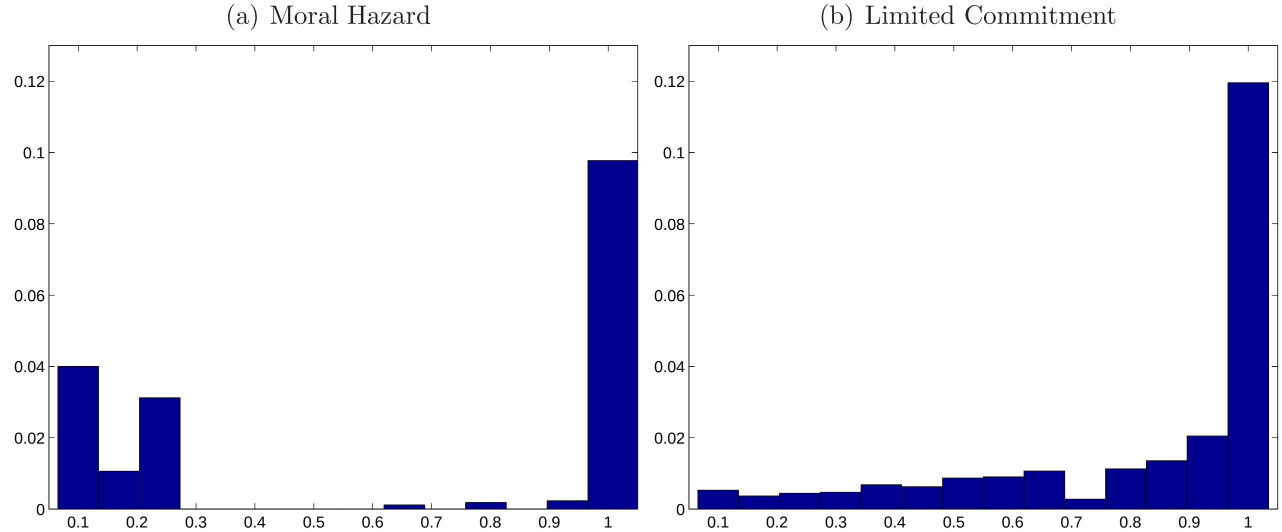
[28]We use an egalitarian welfare measure, namely the sum across all individuals of their present discounted values of utility of consumption with equal Pareto weights (equal to unity)

$$\mathcal{W}^j = \int v_j(a,z) dG_j(a,z),$$

where $j \in \{LC, MH, FB\}$ indexes the specific regime. We use this welfare measure rather than the more standard consumption equivalent measures (Lucas, 2003) because it is considerably easier to compute. In doing so, we follow Buera and Shin (2011). Note also that a possible limitation of our welfare measure is that it does not take into account transition dynamics; it is as if the economy started up at the steady state. This is again for computational reasons; in principle transition dynamics can be easily taken into account. A technical issue arises because with CRRA utility and $\sigma > 1$ individual value functions are negative. In Table 2, "Welfare (% of FB)" in regime $j \in \{LC, MH\}$ is therefore $1 - \frac{\mathcal{W}^{FB} - \mathcal{W}^j}{|\mathcal{W}^{FB}|}$.

[29]The reason that individuals exert lower effort under moral hazard is entirely standard: if intermediaries

Figure 3: Distribution of Entrepreneurial Effort.



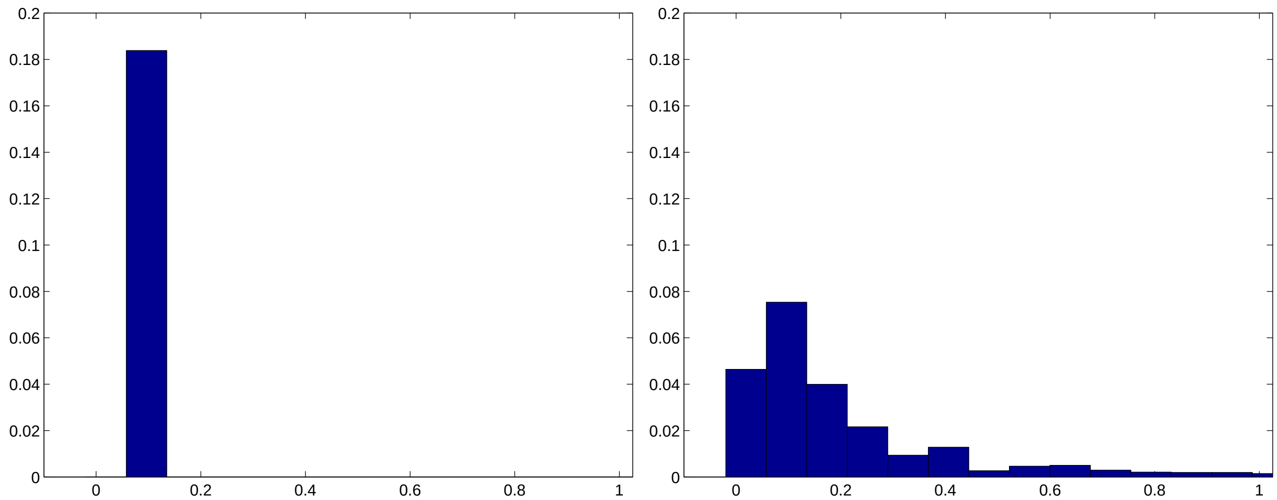
limited commitment (panel b), in contrast, largest single group of entrepreneurs exert the highest possible effort level. The low effort of a large number of entrepreneurs under moral hazard makes it more likely that entrepreneurs draw a low residual productivity realization, thereby depressing firm level TFP. In turn this also depresses aggregate TFP.

Figure 4 plots the distributions of the marginal product of capital in the two regimes. In our moral hazard regime, marginal products of capital are *equalized* across firms so that the distribution of marginal products is degenerate. As already discussed in Section 2.2, this is due to our special formulation of the moral hazard problem in which a change in an entrepreneur's capital stock does not change his incentive to shirk. Since firms do not face any other constraints that limit the amount of capital they can rent, all of them rent capital until their expected

offered the full information contract with full insurance against residual productivity, ε, to individuals, these would always exert low effort. To incentivize individuals, the contract gives up on full insurance and assigns them lower consumption in case of a low realization of residual productivity. But because the optimal contract delivers a given utility level to individuals at the least cost to intermediaries, it trades off providing insurance and incentivizing effort. Implementing full-information effort requires giving up too much insurance and hence having to compensate individuals in some other form. What is not entirely standard is that our analysis takes place in general equilibrium and hence also factor prices change.

Figure 4: Distribution of Marginal Products of Capital.

(a) Moral Hazard (b) Limited Commitment



marginal product equals the user cost of capital:[30]

$$z\bar{\varepsilon}(e)f_k(k,\ell) = r + \delta, \quad \bar{\varepsilon}(e) \equiv \sum_{\varepsilon} \varepsilon p(\varepsilon|e). \tag{13}$$

In contrast, in the limited commitment regime (panel b), the presence of collateral constraints (6) implies that marginal products of capital are not equalized across individual firms, that is capital is misallocated. Of course, more general formulations of the moral hazard problem may result in capital misallocation in addition to depressed effort, but our point is that this is not necessarily the case.[31] In a recent paper Midrigan and Xu (2014) calibrate a model with collateral constraints to the variability and persistence of establishment-level output, capital and employment (but not the dispersion of marginal products), and argue that this model generates fairly small dispersion of marginal products and hence TFP losses from misallocation. Calibrating their model to our model-generated data would lead them to draw the same conclusion. However, the instructive special case of the moral hazard economy just discussed is an example in which there are sizable TFP losses even though returns to capital are equalized

[30] Similarly, entrepreneurs hire labor to equate the expected marginal product of labor to the wage, $z\bar{\varepsilon}(e)f_\ell(k,\ell) = w$. Hence, even though entrepreneurs bear some of the residual productivity risk, ε, under the optimal contract they behave as if they are risk neutral. This is because risk neutral intermediaries find it optimal to first maximize expected profits and to then assign ε-dependent consumption to entrepreneurs to make sure they expend the optimal amount of effort given incentive constraints. Since intermediaries pool risk over a large number of households, the expectation in (13) can be thought of as an integral over the population and not an expectation for the individual.

[31] It is easy to see that if the distribution of residual productivity were some non-degenerate function of capital $p(\varepsilon|e,k)$ rather than $p(\varepsilon|e)$, marginal products of capital would no longer be equalized.

across all firms.[32]

The effect of moral hazard on occupational choose also differs markedly from that of limited commitment. This is shown in Figure 5 which presents occupational choice maps, that is the occupational choice corresponding to different combinations of individual ability and wealth. For the purposes of this paper, the occupational choice effects are of interest mainly as a benchmark for our analysis of the equilibrium interaction of financial frictions present in the same economy in Section 5. Under moral hazard (panel a), selection is mainly on ability, and wealthier individuals are in fact somewhat *less* likely to become entrepreneurs. Two offsetting effects are at work here: a "debt overhang" effect making poorer individuals *less* likely and richer individuals *more* likely to be entrepreneurs (similar to Aghion and Bolton, 1997; Ghosh, Mookherjee and Ray, 2000; Paulson, Townsend and Karaivanov, 2006), and a standard wealth effect on effort supply making richer individuals *less* likely to be entrepreneurs. Under our parameterization, the latter effect dominates, but under alternative parameterizations this result could be overturned. In the frictionless economy in which the "debt overhang" effect is not present the negative relationship between wealth and entrepreneurship is even more pronounced.[33] In contrast, under limited commitment (panel b), selection into entrepreneurship is based on both ability and wealth. Some able but poor individuals cannot rent enough capital to make entrepreneurship attractive. And some less able individuals become entrepreneurs only because they are wealthy.

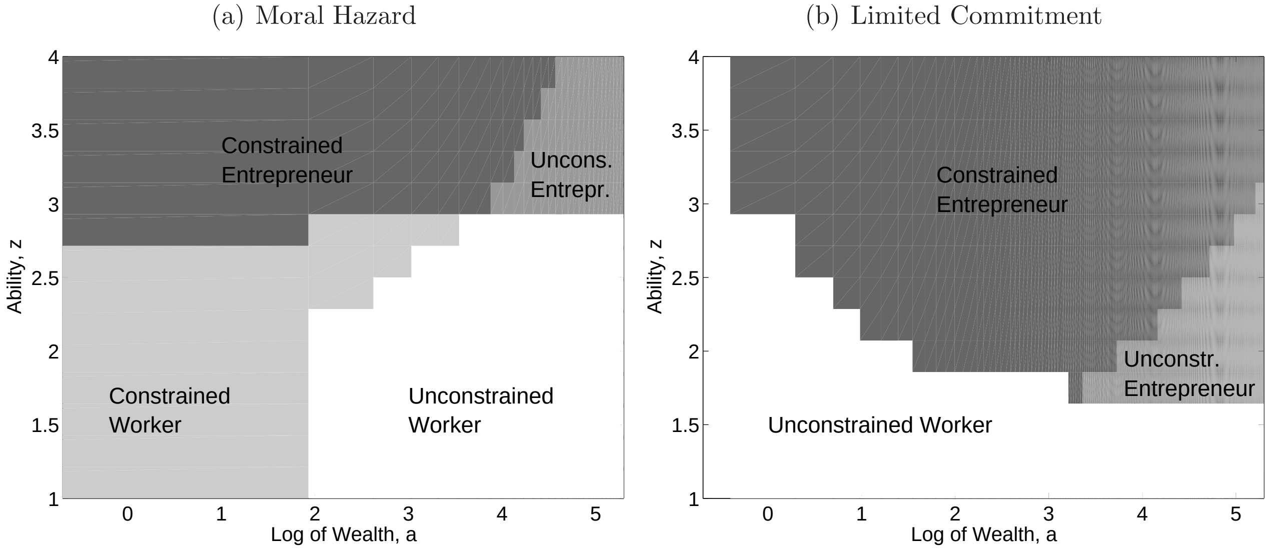
5 Mixtures of Moral Hazard and Limited Commitment

The previous section compared two economies: one in which all agents were subject to the limited commitment friction and another in which all agents were subject to the moral hazard friction. However, there is no reason why a given economy should be subject to only one

[32]In the second part of their paper, Midrigan and Xu examine an alternative mechanism through which financial frictions can lead to large TFP losses in the presence of low dispersion of marginal products: inefficiently low levels of entry and technology adoption. Also see Buera, Kaboski and Shin (2011).

[33]Several remarks are in order. First, the negative wealth effect is stronger for entrepreneurs than for workers because entrepreneurs' profits are more sensitive to effort than workers' labor income (entrepreneur's effort is "leveraged" so to speak). Second, the wealth effect could potentially be eliminated by a different choice of preferences than (11), for example those proposed by Greenwood, Hercowitz and Huffman (1988). See Rampini (2004) for a related analysis of a moral hazard problem with non-separable preferences. However, we find separable preferences with a wealth effect more appealing, not least because they are more standard in dynamic moral hazard problems. Finally, the debt overhang effect in our dynamic model is less pronounced than that in static models like Aghion and Bolton (1997); Ghosh, Mookherjee and Ray (2000); Paulson, Townsend and Karaivanov (2006). In a static model, an entrepreneur's repayments to the intermediary are bounded by his revenues. This gives a strong incentive to shirk to highly indebted individuals. In our dynamic setup, instead, repayments can be postponed to the future through borrowing and lending.

Figure 5: Occupational Choice and Binding Patterns of Constraints



imperfection. For example, a number of studies using micro data from Thailand find that individuals face different types of financial frictions depending on location. We discuss the data and methodology underlying these studies in more detail in section 5.2. Motivated by these findings, we ask in this section: what happens if both frictions are present in the same economy?

5.1 Aggregates and Inter-Sectoral Flow of Funds in Mixture Regimes

To answer this question, we consider economies in which a fraction m of the population are subject to moral hazard and the remaining $1-m$ are subject to limited commitment. We then examine how aggregates vary when we vary from low to high, left to right, as if starting with limited commitment and then putting in some moral hazard m. In this way we can trace out all intermediate cases between the two extremes in the previous section: the pure moral hazard regime, $m=1$, and the pure limited commitment regime, $m=0$. We argue that such *mixture* regimes are different from simple convex combinations of the pure moral hazard and pure limited commitment economies. To be clear, note that we focus on the *equilibrium* interaction of financial frictions rather than the interaction of financial frictions at the *individual level*, i.e. the effect of subjecting a given individual to the two frictions at the same time (see e.g. Paulson, Townsend and Karaivanov, 2006). In principle, our apparatus is flexible enough to also conduct the latter exercise but motivated by the empirical evidence documenting different frictions in different regions within the same economy, we instead focus on the former.

As already mentioned, one possible interpretation is that different financial regimes represent different locations or sectors within an economy: moral hazard in urbanized and industrialized areas and limited commitment in rural and agricultural areas. The parameter m then has the interpretation of the fraction of the population living in urban areas (though our formulation allows workers to live in one location but to work in another, analogous to "temporary migration").

Figure 6 plots various aggregates against the fraction of the population subject to moral hazard, m. Most aggregates in the mixed regime, for example GDP and the interest rate, lie in between their values in the pure limited commitment and pure moral hazard economies. But this is not true for all variables. Notably, TFP is a non-monotone function of m. As we discuss in more detail at the end of this section, such non-monotonicities arise from general equilibrium effects and would not have been obtained in a partial equilibrium with a fixed interest rate.

To better understand these mixed regimes, we consider in more detail the economy in which fifty percent of the population are subject to moral hazard and the remaining fifty percent are subject to limited commitment, that is $m = 0.5$. Column 1 of Table 3 reports some aggregate statistics of the economy with both regimes. Again, TFP is *higher* in the mixed regime than in either of the two pure regimes. To understand where this comes from, columns 2 and 3 of Table 3 report separate statistics for the limited commitment and moral hazard sectors of the same mixed-regime economy. Consider first TFP. Column 2 shows that sectoral TFP

Table 3: Comparison of LC and MH Sectors within Mixed Regime

	Mixed Regime, m=0.5	MH sector	LC sector
	(a) National and Sectoral Aggregates		
TFP (% of FB)	0.895	0.819	1.092
GDP (% of FB)	0.855	1.051	0.660
Welfare (% of FB)	0.830	0.875	0.786
Wage (%of FB)	0.887	0.887	0.887
Interest Rate	0.003	0.003	0.003
% Entrepreneurs	0.176	0.176	0.175
	(b) Importance of Sectors in Aggregate Economy		
% Sector Contributes to GDP		0.614	0.386
% of Labor Employed in Sector		0.615	0.386
% of Capital Used in Sector		0.757	0.243
% of Labor Supplied by Sector		0.471	0.529
% of Capital Supplied by Sector		0.534	0.466
	(c) Intersectoral Capital and Labor Flows		
Net Labor Inflow (% of used)		0.231	-0.377
Net Capital Inflow (% of used)		0.302	-0.894

in the LC sector is 109 percent of first-best TFP, that is much higher than TFP in the pure limited commitment regime (column 1 of Table 2). To understand this, consider Figure 7 which presents the occupational choice maps for the two sectors in the mixed regime. Compare the

Figure 6: Aggregate Impact of Importance of Moral Hazard vs. Limited Commitment, m

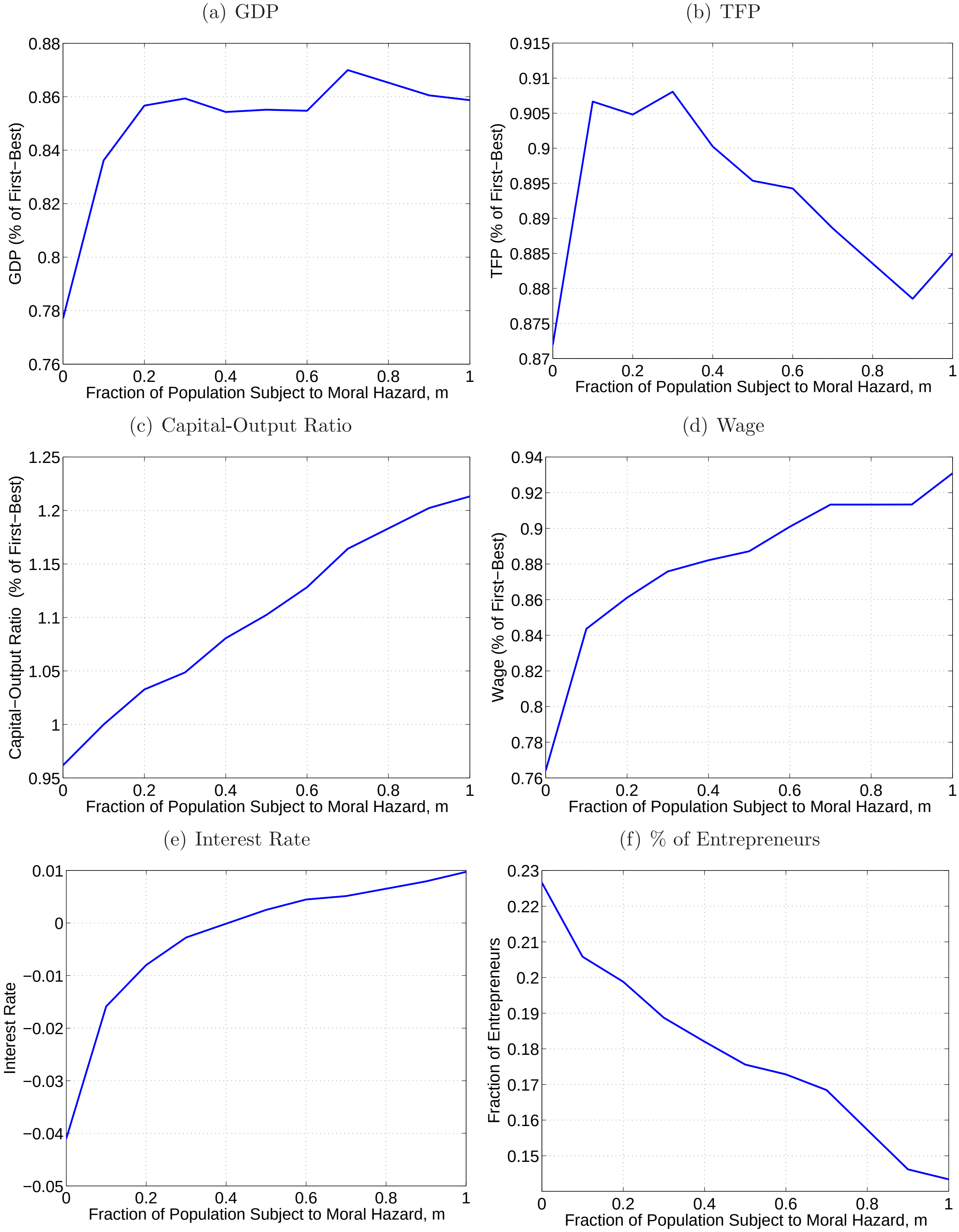
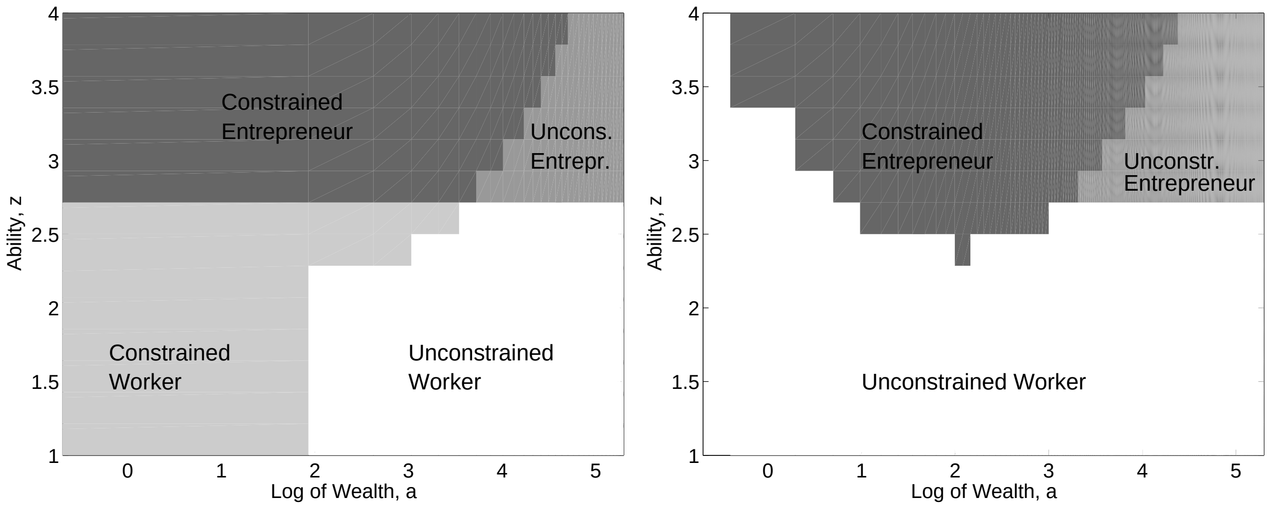


Figure 7: Occupational Choice in LC and MH sectors of Mixed Regime

(a) Moral Hazard (b) Limited Commitment



map for limited commitment regime in panel (b) of Figure 7 to its analogue for the pure limited commitment regime in panel (a) of Figure 5. Strikingly, the LC sector in the mixed regime does not feature the very low ability entrepreneurs of the pure LC economy. This is because the relatively high interest rate in the mixed regime makes it unattractive for untalented individuals to become entrepreneurs as their revenues are too low to cover their cost of capital. This results in high sectoral and aggregate TFP. The high wage in the mixed regime relative to column 2 of Table 2 is then a direct consequence of high aggregate TFP.

Panel (b) of Table 3 shows the fraction of labor and capital used and supplied by the two sectors. The MH sector uses more than three quarters of the economy's capital. This is because the interest rate in the mixed regime lies in between those in the pure limited commitment and moral hazard economies, making capital relatively cheap for the MH sector but relatively expensive for the LC sector. Finally, the MH sector supplies less than half of the economy's labor but more than half its capital. Panel (c) of Table 3 presents the same information in another way, by computing intersectoral capital and labor flows. The MH sector experiences both capital and labor inflows. For example, 30.2 percent of the capital used in the MH sector is owned by the LC sector and hence the MH sector is a net borrower (and conversely the LC sector is a net lender). We also compare welfare in the two sectors of the mixed regime to welfare in the pure moral hazard and limited commitment regimes, corresponding to a comparison of steady state welfare before and after integration of the two sectors (with no transition). Interestingly, we find that while integration always increases welfare in the LC sector, welfare in the MH sector can decrease.[34]

[34] These are numerical results. In particular, we computed welfare for all alternative parameter values we consider in Section 6. In all of them financial integration leads to an increase in welfare in the LC sector. In

5.2 Regional Variation of Financial Frictions

As noted earlier, various papers have studied within-country regional variation in the type of financial friction faced by individuals. This section first briefly summarizes these studies and then discusses implications of these findings for our theory. Most of the studies use Townsend Thai data, a rich micro dataset from Thailand.[35] Using these data and a variety of approaches, these studies have come to the conclusion that moral hazard is a better fit in urban areas and in the more developed Central region, while limited commitment is a better fit in rural areas and in the less developed Northeast region. In Karaivanov and Townsend (2014), the principle determining factor is the persistence of the capital stock; capital stock levels are quite persistent, that is move slowly, despite the potential for funds to flow in to allow quicker adjustment. In contrast, ignoring information on investment and using consumption and income data only in these rural areas, one gets an alternative picture, that is, risk sharing seems quite good and an endogenous mechanism design regime appears to fits the data better. But again, to make the point, even in these rural areas, Samphantharak and Townsend (2009) and Kinnan and Townsend (2012) find that investment, unlike consumption, is sensitive to cash flow and that the presence of kinship networks only partially mitigates this conclusion.

Paulson, Townsend and Karaivanov (2006) fit an occupation choice model, mapping wealth into transitions from agriculture and wage earning categories into enterprise; not only is this quantitative occupation choice mapping consistent with regional findings, moral hazard in the Central region and mixed regimes in the Northeast, so are the financial contracts in the sense that debt to asset ratios are decreasing with wealth in the Central region, consistent with moral hazard, as self financing mitigates the information obstacles or debt overhang, and not in the Northeast (under limited commitment, constrained businesses would increase investment as wealth increases).

Finally, Ahlin and Townsend (2007) use variables to capture the signaling, monitoring of project choice, screening, and informal penalties that are key in various models of joint liability

all but two of the our parameterizations, welfare *decreases* in the MH region after financial integration. The two exceptions are extreme values for the Frisch elasticity φ and the persistence of talent ρ (which one would expect to matter for the severity of the moral hazard problem). Results are available upon request.

[35] The baseline survey of the Townsend Thai data was conducted in 1997 covering four provinces, two in the relatively wealthy and industrializing Central region near Bangkok, Chacheongsao and Lopburi, and two in the relatively poor, semi-arid Northeast, Buriram and Srisaket, with 192 villages overall, 15 households per village or 2880 household overall. An annual resurvey on approximately one third of this sample began one year later and has continued up to the present, 16 years. A monthly household-level survey data gathered from 16 villages in the original baseline began in August 1998 now spans 192 months and has detailed information on the financial accounts of households. Samphantharak and Townsend (2009) provide a more detailed description of this. The Townsend Thai survey was further expanded to more provinces in the North and South of the country in 2003, 2004 and to Urban communities in 2005. We use sub-periods of the rural and urban data in a comparison below.

(Stiglitz, 1990; Banerjee, Besley and Guinnane, 1994; Ghatak, 1999; Besley and Coate, 1995) and find that information obstacles are salient in the more urban and industrialized Central region and that limited commitment is an issue in the more urban and agricultural Northeast.

When we take these findings at face value and interpret the MH sector as urban and industrialized areas and the LC sector as rural and agricultural areas, our economy produces interregional patterns of aggregate income, capital and labor flows and external finance that resemble rural-urban patterns observed in the data. In particular, regional income, the aggregate capital stock and the amount of external finance are higher in urban areas, and both capital and labor flow from rural to urban areas.[36]

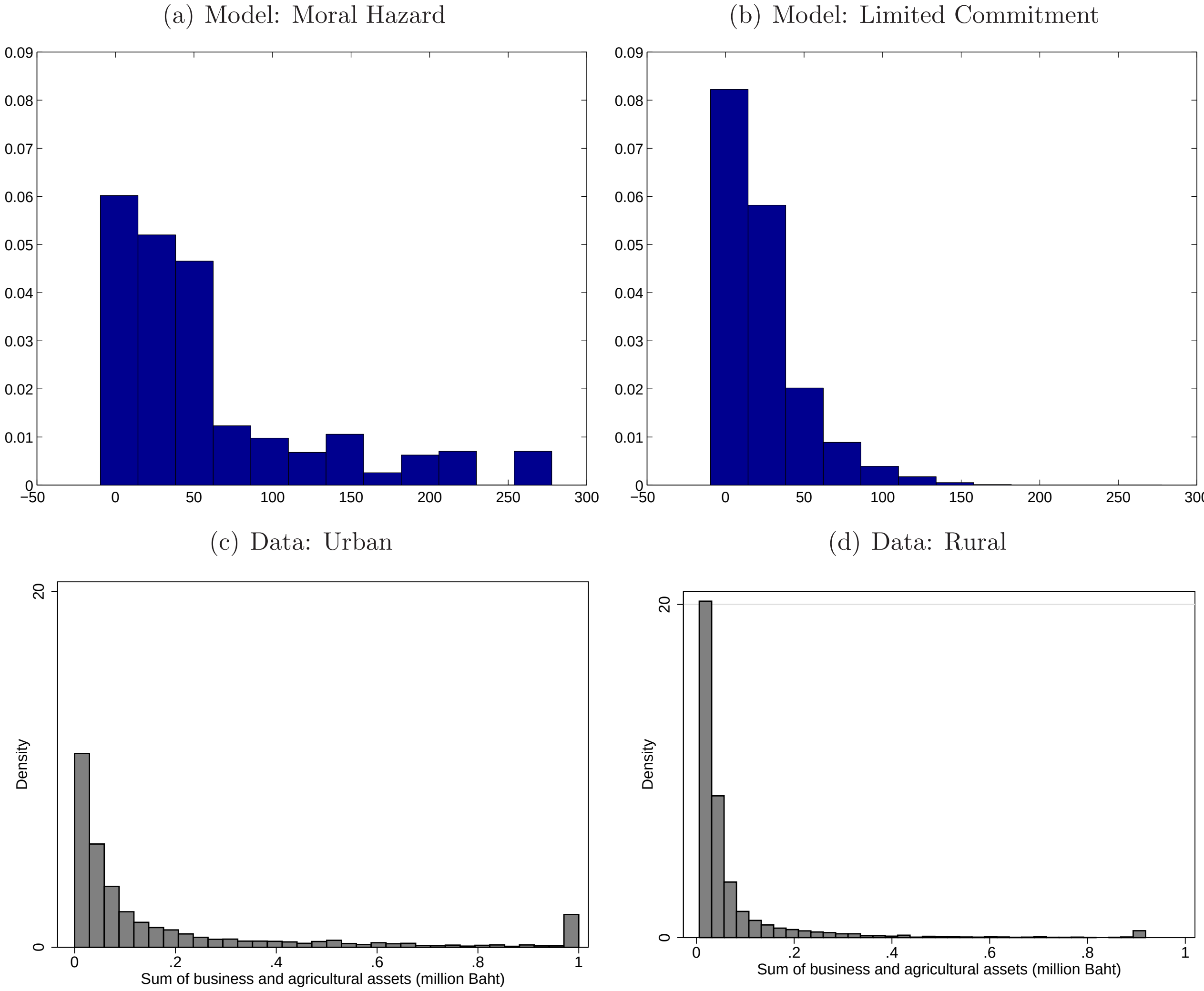
As already discussed in section 4, our model also predicts that the distributions of various micro-level variables, such as firm size, differ across sectors. It is therefore natural to ask how the differences in such distributions between the LC and MH sectors in the model compare to those observed between rural and urban areas in the data, though the data were not used in this way previously in the other studies cited. Panels (a) and (b) of Figure 8 plots the distribution of firm size as measured by a firm's productive assets (its capital) in the LC and MH sectors. Panels (c) and (d) plot the same objects using data from Thailand (Townsend Thai data) described in more detail in footnote 35.[37] The distributions in our model and the data share some remarkable similarity. First, the firm size (asset) distribution is much more compressed in rural areas and in the LC sector. Second, the firm size distribution in the model's MH sector does have a large number of small firms (the first three bins are large) but then it has a relatively fat right tail. So does the size distribution in urban areas.

Of course, in practice there are many other factors that distinguish cities from villages and industrialized from agricultural areas (for example, cities have better infrastructure, higher population density, regions vary in resource base etc). But we nevertheless find it noteworthy

[36]Our preferred interpretation of the labor flows from rural to urban areas is as temporary migration which is a particularly wide-spread phenomenon in developing countries (see e.g. Morten, 2013). This interpretation is consistent with our assumption that individuals are subject to the financial regime of their region of origin rather than their workplace (e.g. individuals from the LC sector (rural area) are subject to limited commitment and perfect risk-sharing of residual productivity even though they work in the MH sector (city)). An interesting extension would be to examine the feedback from temporary migration to participation in risk-sharing arrangements back in the village as in Morten (2013).

[37]The plots use the 2005-2011 waves of the Townsend Thai Data from four provinces (Lopburi, Chachoengsao, Buriram, and Sisaket). The urban dataset covers 96 urban communities and is a stratified, clustered, random sample of 15 households in each community. The rural sample is collected over the same time period and covers four villages in each province and about forty households per village. Firm size is defined as the sum of agricultural and business assets, and we drop households who report zero holdings of either category, leaving us with 601 urban and 659 rural households. We chose assets as a measure of a firm's size rather than employment (as is perhaps more standard), because of the prevalence of self-employed individuals (i.e. few paid employees) in the Townsend Thai data. For comparison with the rural data, the urban data are winsorized at 1 million baht.

Figure 8: Firm Size (Capital) Distribution in LC and MH sectors of Mixed Regime

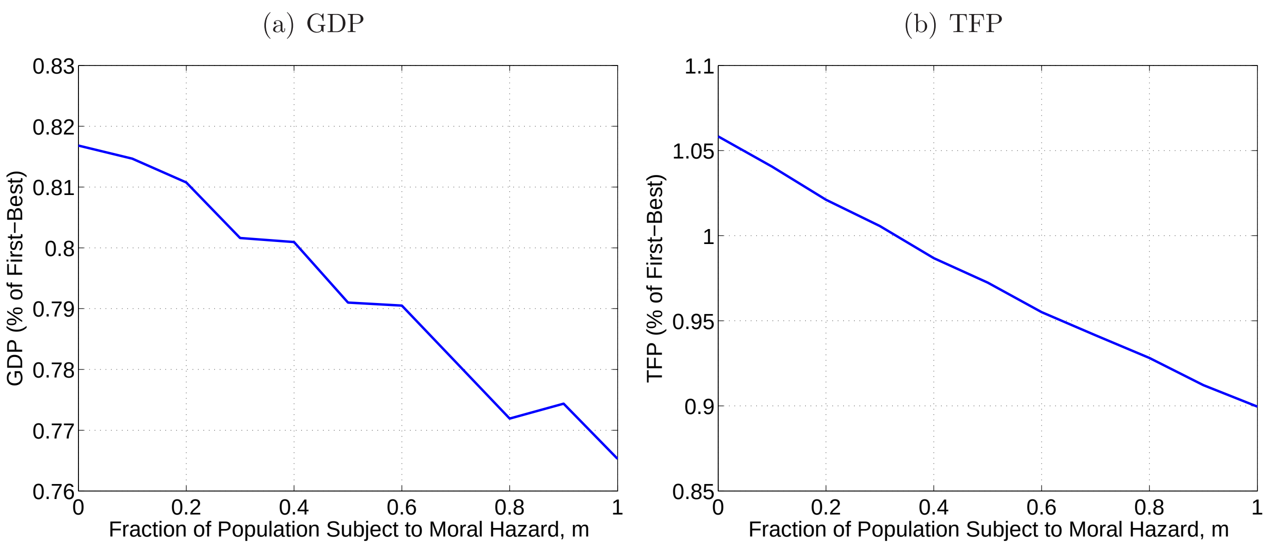


that we can generate a number of observed rural-urban patterns by letting only the financial regime differ across these regions. Evidently that is playing a contributory, potentially large role.

5.3 Importance of General Equilibrium Effects

Finally, note that the non-monotonicity in TFP that we have uncovered in Figure 6 is due to general equilibrium effects. To show this, Figure 9 presents results for the same economy but for the case where the interest rate is fixed at the value of the frictionless economy.[38] In contrast to their counterparts in general equilibrium (Figure 6), in partial equilibrium all variables are monotonic functions of m. Similarly some of our results about capital and labor flows between

Figure 9: Aggregate Impact of Varying m in Partial Equilibrium



the MH and LC sectors change in partial equilibrium. To this end, Table 4 replicates Table 3 but in partial equilibrium. Focussing on panel (c) and comparing it to panel (c) of Table 3 it can be seen that in partial equilibrium *both* the MH and LC sector experience capital outflows. This is because our partial equilibrium is computed at the frictionless interest rate which is strictly higher than the mixed regime interest rate in a closed economy, that is there is a "world interest rate" which is higher than the return on savings that can be obtained in either the

[38]We have also computed equilibria at both the frictionless interest rate and wage. These are somewhat hard to interpret because the wage is so elevated that hardly any of the households choose to be entrepreneurs. We therefore prefer to let the wage adjust so as to clear the labor market, and to compare an economy with only the interest rate fixed with one in which the interest rate (and the wage) are determined in equilibrium. The interest rate is also the price that is more obviously tied to differences in financial regimes because it affects borrowing and lending decisions, and the interest rate is typically taken as given in small open economy papers in the literature whereas the wage is not.

Table 4: Comparison of LC and MH Sectors within Mixed Regime in Partial Equilibrium

	Mixed Regime, m=0.5	MH sector	LC sector
	(a) National and Sectoral Aggregates		
TFP (% of FB)	0.972	0.905	1.048
GDP (% of FB)	0.791	0.741	0.841
Welfare (% of FB)	0.972	0.960	0.984
Wage (%of FB)	0.838	0.838	0.838
Interest Rate	0.025	0.025	0.025
% Entrepreneurs	0.160	0.138	0.181
	(b) Importance of Sectors in Aggregate Economy		
% Sector Contributes to GDP		0.469	0.531
% of Labor Employed in Sector		0.468	0.532
% of Capital Used in Sector		0.554	0.446
% of Labor Supplied by Sector		0.492	0.508
% of Capital Supplied by Sector		0.474	0.526
	(c) Intersectoral Capital and Labor Flows		
Net Labor Inflow (% of used)		-0.054	0.043
Net Capital Inflow (% of used)		-1.193	-2.026

LC or MH regimes. That being said, relative to Table 3 it remains true that the LC sector experiences the larger capital outflows than the MH sector and therefore the results are as expected. Much more surprisingly, relative to Table 3, the direction of labor flows is reversed: now labor flows from the MH sector to the LC sector (though the magnitudes are small).

6 Robustness of Main Results

6.1 Robustness to Alternative Paramaterizations

The numerical results we have presented were computed with the parameter values in Table 1. In this section, we briefly discuss in how far each of our results is robust to alternative parameterizations. We have computed results under fourteen alternative parameterizations. [39] We here only briefly summarize these with an emphasis on differences in results, but all detailed results are available on request.

Aggregates in moral hazard and limited commitment regimes (Table 2). First, the equilibrium interest rate is of similar magnitude in all our alternative parameterizations and always considerably smaller under limited commitment than under moral hazard. In almost all our alternative parameterizations, TFP and GDP losses are of similar magnitude (within five

[39] We have studied variations in the parameters $\beta, \varphi, \chi, \rho, \theta$ and λ.

percentage points) as those reported in Table 2. Not surprisingly, one exception are variations in the parameter λ governing the tightness of collateral constraints but the differences are in line with results in the existing literature (Buera and Shin, 2013; Buera, Kaboski and Shin, 2011; Midrigan and Xu, 2014).[40] A second exception is the parameterization with lower disutility of labor $\chi = 0.1$ which results in considerably larger GDP losses (but without changing TFP losses by much).[41] A third exception is TFP in the moral hazard regime when we decrease the persistence of entrepreneurial talent from $\rho = 0.75$ all the way to $\rho = 0$ which results in TFP losses of only two percent. The wage and the percentage of entrepreneurs are more sensitive to parameter values, which is why we do not emphasize these results as much.

Behavior of aggregates when varying fraction of population subject to moral hazard, m (Figure 6). In all our alternative parameterizations, aggregates in mixture regimes are different from simple convex combinations of the pure limited commit and pure moral hazard regimes (that is, the variables plotted in Figure 6 are not simply linear functions of m). In contrast but not surprisingly, the related result that aggregate TFP is a non-monotonic function of the prevalence of moral hazard m (see panel (b) of Figure 6) is much more sensitive to parameter values. For example, if $\lambda = 1$ instead of $\lambda = 1.8$ as in our benchmark exercise, TFP in the pure limited commitment regime ($m = 0$) is much lower. But TFP in the pure moral hazard regime ($m = 1$) does not depend on λ. The difference between the two extremes is then so large that TFP is no longer non-monotonic for intermediate values of m.

Inter-sectoral (regional) patterns of income, external finance, and capital and labor flows (Table 3). These results are very robust across all our alternative parameterizations. In particular, sectoral (regional) income, the aggregate capital stock, and the amount of external finance are always higher in the MH sector (urban and industrialized areas), and both capital and labor always flow from the LC sector (rural and agricultural areas) to the MH sector (urban and industrialized areas).

6.2 Robustness to Alternative Functional Forms

Our numerical results were computed using the functional forms for the utility and production functions, and the talent and residual productivity distributions described in section 3. It is well-known that in moral hazard problems, the functional form of the utility function can be

[40]Decreasing λ from 1.8 to 1 results in higher TFP losses of 28.3 percent and increasing λ to 3 results in smaller TFP losses of 9 percent.

[41]This is because setting $\chi = 0.1$ rather than $\chi = 0.525$ results in much larger first-best labor supply. Therefore also first-best GDP, i.e. the denominator in calculations of GDP losses is much larger with lower disutility of labor.

important, in particular whether it is separable. To explore this, we have also computed results for the case where the utility function takes the non-separable form proposed by Greenwood, Hercowitz and Huffman (1988), i.e. there is no wealth effect. This matters for some results but not for others. For example, as is well known, the inverse Euler equation (Lemma 1 in Appendix E.1) does not hold anymore. Similarly, the occupational choice patterns in the MH regime in Figure 5 are now different because there is no longer a wealth effect making rich individuals less likely to exert effort and hence to be entrepreneurs. It should also be relatively easy to compute results for alternative (say CES) production functions, and talent and residual productivity distributions, but we do not have any strong reasons to believe that these would yield different results.

7 Conclusion

More research is needed that makes use of micro data and takes seriously the micro financial underpinnings of macro models. One likely reason for the relative scarcity of such studies is the lack of reliable balance sheet data for firms, including smaller household-firms and household wage earners in developing countries. The Townsend Thai data used by many of the studies that motivated this paper are an exception. See in particular Samphantharak and Townsend (2009) who construct household balance sheets by treating them in the same way accountants treat corporate firms. The collection of more such data is a very worthwhile project and for Thailand there are plans to collect such data countrywide. An obvious recommendation is also for other countries to collect such balance sheet data. Some of this could be done as part of the collection of manufacturing censuses that many developing are now collecting quite effectively.

In addition to more and better data, we also need more theoretical research in macroeconomics aimed at furthering our understanding of heterogeneous agent models and the complexities these may entail, especially when more realistic micro financial underpinnings and elements from contract theory are incorporated as in the present paper. In our framework, for example, a financial reform affecting underlying obstacles in one of our financial regimes or regions will set in motion intricate transition dynamics of reallocation both across firms and across regions, and of aggregates like GDP and TFP. In ongoing work, we are currently exploring these possibilities and such transitions do seem to be computable. But much more generally, there should be high payoffs from a better theoretical understanding of heterogeneous agent models and the development of better numerical methods needed to compute them.

Our bottom line: not only does the financial sector matter for real variables, including growth and inequality, but also the details of financial contracts matter for the macro economy. This joins what have been largely two distinct literatures – macro development and micro development – into a coherent whole. The macro development literature needs to take into account the contracts we see on the ground and the micro development literature needs to take into account general equilibrium effects of interventions.

Appendix

A Proof of Lemma 1

The Lagrangean for (3) to (5) is

$$\mathcal{L} = \sum_{\varepsilon} p(\varepsilon|e)\left\{U(c(\varepsilon)) - V(e) + \beta \mathbb{E}_{z'} v[a'(\varepsilon), z']\right\}$$

$$+\psi\left[(1+r)a + \sum_{\varepsilon} p(\varepsilon|e)\left\{x[z\varepsilon f(k,\ell) - w\ell - (r+\delta)k] + (1-x)w\varepsilon\right\} - \sum_{\varepsilon} p(\varepsilon|e)\left\{c(\varepsilon) + a'(\varepsilon)\right\}\right]$$

$$+\sum_{e,\hat{e},x} \mu(e,\hat{e},x)\left[\sum_{\varepsilon} p(\varepsilon|e)\left\{U(c(\varepsilon)) - V(e) + \beta \mathbb{E}_{z'} v[a'(\varepsilon), z']\right\} - \sum_{\varepsilon} p(\varepsilon|\hat{e})\left\{U(c(\varepsilon)) - V(\hat{e}) + \beta \mathbb{E}_{z'} v[a'(\varepsilon), z']\right\}\right]$$

The first-order conditions with respect to $c(\varepsilon)$ and $a'(\varepsilon)$ are

$$\psi p(\varepsilon|e) = p(\varepsilon|e)U'(c(\varepsilon)) + \sum_{e,\hat{e},x} \mu(e,\hat{e},x)[p(\varepsilon|e) - p(\varepsilon|\hat{e})]U'(c(\varepsilon))$$

$$\psi p(\varepsilon|e) = p(\varepsilon|e)\beta \mathbb{E}_{z'} v_a(a'(\varepsilon), z') + \sum_{e,\hat{e},x} \mu(e,\hat{e},x)[p(\varepsilon|e) - p(\varepsilon|\hat{e})]\beta \mathbb{E}_{z'} v_a(a'(\varepsilon), z')$$

Rearranging

$$\frac{p(\varepsilon|e)}{U'(c(\varepsilon))} = \frac{1}{\psi}\left[p(\varepsilon|e) + \sum_{e,\hat{e},x} \mu(e,\hat{e},x)[p(\varepsilon|e) - p(\varepsilon|\hat{e})]\right] \tag{14}$$

$$\frac{p(\varepsilon|e)}{\beta \mathbb{E}_{z'} v_a(a'(\varepsilon), z')} = \frac{1}{\psi}\left[p(\varepsilon|e) + \sum_{e,\hat{e},x} \mu(e,\hat{e},x)[p(\varepsilon|e) - p(\varepsilon|\hat{e})]\right] \tag{15}$$

Summing (14) over ε,

$$\sum_{\varepsilon} \frac{p(\varepsilon|e)}{U'(c(\varepsilon))} = \frac{1}{\psi}$$

The envelope condition is

$$v_a(a,z) = \psi(1+r) = (1+r)\left(\sum_{\varepsilon} \frac{p(\varepsilon|e)}{U'(c(\varepsilon))}\right)^{-1} \tag{16}$$

From (14) and (15)

$$U'(c(\varepsilon)) = \beta \mathbb{E}_{z'} v_a(a'(\varepsilon), z') \tag{17}$$

Combining (16) and (17) yields (30).□

B Capital Accumulation

The purpose of this section is to spell out in detail how capital accumulation works in our economy. We assume that there is a representative capital producing firm that holds bonds, B_t, issues dividends, D_t, invests, I_t, to accumulate capital, K_t which it rents out to households at a rental rate R_t. The budget constraint of the capital producer is then

$$B_{t+1} + I_t + D_t = R_t K_t + (1 + r_t)B_t, \quad K_{t+1} = I_t + (1 - \delta)K_t$$

The entire debt of the representative capital producer is held by intermediaries that contract with individuals and hold their wealth, a. Hence the debt market clearing condition is

$$B_t + \int a dG_t(a, z) = 0, \quad \text{all } t. \tag{18}$$

The capital producer maximizes

$$V_0 = \sum_{t=0}^{\infty} \frac{D_t}{\prod_{s=0}^{t}(1 + r_s)}.$$

subject to

$$K_{t+1} + B_{t+1} + D_t = (R_t + 1 - \delta)K_t + (1 + r_t)B_t \tag{19}$$

It is easy to show that this maximization implies the no arbitrage condition $R_t = r_t + \delta$.[42] Therefore the budget constraint (19) is

$$D_t = (1 + r_t)(K_t + B_t) - K_{t+1} - B_{t+1}$$

and so the present value of profits is

$$V_t = \sum_{s=0}^{\infty} \frac{D_{t+s}}{\prod_{\tau=0}^{s}(1 + r_{t+\tau})} = (1 + r_t)(K_t + B_t) \quad \text{all } t.$$

Zero profits implies $K_t + B_t = 0$ for all t. Using bond market clearing (18), this implies that the economy's aggregate capital stock equals its total wealth

$$K_t = \int a dG_t(a, z), \quad \text{all } t.$$

[42] Defining cash-on-hand, $M_t = (R_t + 1 - \delta)K_t + (1 + r_t)B_t$, the associated dynamic program is

$$V_t(M) = \max_{K',B'} \ M - K' - B' + (1 + r_t)^{-1} V_{t+1}[(R_{t+1} + 1 - \delta)K' + (1 + r_{t+1})B']$$

The first order conditions imply $R_{t+1} = r_{t+1} + \delta$.

C Connection of Private Information Regime to Optimal Dynamic Contract

We here show how the our formulation of the contracting problem under moral hazard, (3) to (5), is related to a more familiar formulation of an optimal dynamic contracting problem under private information. In particular, we show that there is optimal insurance against residual productivity shocks, ε, (in a sense defined precisely momentarily) but no insurance against ability shocks, z. We show that for the special case in which there are only residual productivity shocks and ability is deterministic,[43] our formulation is equivalent to an optimal dynamic contracting problem. That is, there is optimal insurance against residual productivity shocks (subject to incentive compatibility) in this special case. The more general formulation (3) to (5) is then simply this special case with uninsurable ability shocks "added on top".

C.1 Equivalence for Special Case with only Residual Productivity (ε) but no Ability (z) Shocks

Standard Formulation with Promised Utility. Consider the following problem: maximize intermediary profits (the PDV of income, y_t given by (2), minus consumption transfers to the agent, c_t)

$$\Pi_t = \mathbb{E}_t \sum_{\tau=t}^{\infty} \frac{y_\tau - c_\tau}{\prod_{s=t}^{\tau}(1+r_s)}$$

subject to providing promised utility of at least W_t to the household

$$\mathbb{E}_t \sum_{\tau=t}^{\infty} \beta^{\tau-t} u(c_\tau, e_\tau) \geq W_t$$

and an incentive compatibility constraint for the household. Assume that there are only residual productivity shocks (ε) and that entrepreneurial ability (z) is deterministic and fixed over time. Without loss of generality, set $z = 1$. To simplify notation, define by $Y(\varepsilon, e)$ an household's income given optimal choices for capital, labor and occupation

$$Y(\varepsilon, e) = \max_{x,k,\ell} \{x[\varepsilon f(k,\ell) - w\ell - (r+\delta)k] + (1-x)w\varepsilon\}.$$

[43] That is, the transition probabilities for entrepreneurial talent are degenerate, $\mu(z'|z) = 1$ if $z' = z$ and zero otherwise.

If $W_t = W$ is promised to the household, the intermediary's value $\Pi_t = \Pi(W_t)$ satisfies the Bellman equation

$$\begin{aligned}
&\Pi(W) = \max_{e,c(\varepsilon),W'(\varepsilon)} \sum_{\varepsilon} p(\varepsilon|e)\left\{Y(\varepsilon,e) - c(\varepsilon) + (1+r)^{-1}\Pi[W'(\varepsilon)]\right\} \quad \text{s.t.} \\
&\sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e] + \beta W'(\varepsilon)\right\} \geq \sum_{\varepsilon} p(\varepsilon|\hat{e})\left\{u[c(\varepsilon),\hat{e}] + \beta W'(\varepsilon)\right\} \;\; \forall e,\hat{e} \\
&\sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e] + \beta W'(\varepsilon)\right\} = W.
\end{aligned} \tag{P1}$$

where we have used that the stream of household income is (2).

Equivalence: The joint budget constraint of a risk-sharing syndicate is

$$a_{t+1} = y_t - c_t + (1+r_t)a_t.$$

This can be written in present-value form as

$$0 = \pi_t + a_t(1+r), \quad \text{for all } t \quad \text{where} \quad \pi_t \equiv \mathbb{E}_t \sum_{\tau=t}^{\infty} \frac{y_\tau - c_\tau}{\prod_{s=t}^{\tau}(1+r_s)} \tag{20}$$

are the intermediary's expected future profits. Equivalently

$$\pi_t + \mathbb{E}_t \sum_{\tau=t}^{\infty} \frac{c_\tau}{\prod_{s=t}^{\tau}(1+r_s)} = \mathbb{E}_t \sum_{\tau=t}^{\infty} \frac{y_\tau}{\prod_{s=t}^{\tau}(1+r_s)} + (1+r_t)a_t.$$

which says that the intermediary's expected profits plus the expected present value of future consumption must equal total income of a risk-sharing syndicate. We can use (20) to establish a useful equivalence result.

Proposition 1 *Suppose the Pareto frontier* $\Pi(W)$ *is decreasing at all values of promised utility,* W*, that are used as continuation values at some point in time. Then the following dynamic program is equivalent to* (P1)

$$\begin{aligned}
&v(a) = \max_{e,c(\varepsilon),a'(\varepsilon)} \sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e] + \beta v[a'(\varepsilon)]\right\} \quad s.t. \\
&\sum_{\varepsilon} p(\varepsilon|e)\left\{u[c(\varepsilon),e] + \beta v[a'(\varepsilon)]\right\} \geq \sum_{\varepsilon} p(\varepsilon|\hat{e})\left\{u[c(\varepsilon),\hat{e}] + \beta v[a'(\varepsilon)]\right\} \;\; \forall e,\hat{e} \\
&\sum_{\varepsilon} p(\varepsilon|e)\left\{c(\varepsilon) + a'(\varepsilon)\right\} = \sum_{\varepsilon} p(\varepsilon|e)Y(\varepsilon,e) + (1+r)a
\end{aligned} \tag{P2}$$

Proof: The proof has two steps.

Step 1: write down dual to (P1)**.** Because the Pareto frontier $\Pi(W)$ is decreasing at the W under consideration, we can write the last constraint of (P1) (promise-keeping) with a (weak) inequality rather than an inequality. This does not change the allocation chosen under the optimal contract.[44] The dual to (P1) is then to maximize

$$V(\pi)=\max_{e,c(\varepsilon),\pi'(\varepsilon)}\sum_{\varepsilon}p(\varepsilon|e)\left\{u[c(\varepsilon),e]+\beta V[\pi'(\varepsilon)]\right\}\quad\text{s.t.}$$
$$\sum_{\varepsilon}p(\varepsilon|e)\left\{u[c(\varepsilon),e]+\beta V[\pi'(\varepsilon)]\right\}\geq\sum_{\varepsilon}p(\varepsilon|\hat{e})\left\{u[c(\varepsilon),\hat{e}]+\beta V[\pi'(\varepsilon)]\right\}\ \forall e,\hat{e}\tag{P1'}$$
$$\sum_{\varepsilon}p(\varepsilon|e)\left\{Y(\varepsilon,e)-c(\varepsilon)+(1+r)^{-1}\pi'(\varepsilon)\right\}\geq\pi.$$

where $\pi=\Pi(W)$. Because $\Pi(W)$ is decreasing, its inverse $V(\pi)$ is also decreasing. We can therefore replace the inequality in the last constraint of (P1') with an equality.

Step 2: express dual in terms of asset position rather than profits. Let

$$\pi=-a(1+r),\quad \pi'(\varepsilon)=-a'(\varepsilon)(1+r).\tag{21}$$

Substituting (21) into (P1') and defining $v(a)=V[-(1+r)a]$, yields (P2).□

The change of variables (21) simply uses the present-value budget constraint (20) to express the problem in terms of assets rather than the PDV of intermediary profits.

C.2 General Case: Comparison of Our Formulation with Optimal Contract

Optimal Contracting Problem. Consider the following problem: maximize intermediary profits

$$\Pi_t=\mathbb{E}_t\sum_{\tau=t}^{\infty}\frac{y_\tau-c_\tau}{\prod_{s=t}^{\tau}(1+r_s)}$$

subject to providing promised utility of at least W_t to the household

$$\mathbb{E}_t\sum_{\tau=t}^{\infty}\beta^{\tau-t}u(c_\tau,e_\tau)\geq W_t$$

and an incentive compatibility constraint for the household. If $W_t=W$ is promised to the household and its current ability shock is $z_t=z$, the intermediary's value $\Pi_t=\Pi(W_t,z_t)$ satisfies the Bellman equation

$$\Pi(W,z)=\max_{e,c(\varepsilon),W'(\varepsilon)}\sum_{\varepsilon}p(\varepsilon|e)\left\{Y(\varepsilon,z,e)-c(\varepsilon)+(1+r)^{-1}\mathbb{E}_{z'}\Pi[W'(\varepsilon),z']\right\}\quad\text{s.t.}$$
$$\sum_{\varepsilon}p(\varepsilon|e)\left\{u[c(\varepsilon),e]+\beta W'(\varepsilon)\right\}\geq\sum_{\varepsilon}p(\varepsilon|\hat{e})\left\{u[c(\varepsilon),\hat{e}]+\beta W'(\varepsilon)\right\}\ \forall e,\hat{e}\tag{P3}$$
$$\sum_{\varepsilon}p(\varepsilon|e)\left\{u[c(\varepsilon),e]+\beta W'(\varepsilon)\right\}=W.$$

[44] Note that this would not be the case if $\Pi(W)$ would be increasing. In that case, replacing the equality by an inequality would change the allocation because it would deliver strictly higher welfare to both parties.

where

$$Y(\varepsilon, z, e) = \max_{x,k,\ell} \{x[z\varepsilon f(k,\ell) - w\ell - (r+\delta)k] + (1-x)w\varepsilon\}$$

Compare this formulation to the one used in the main text, (3) –(5). Note that under the optimal contract (P3), utility $W(\varepsilon)$ cannot depend on z'. That is, the principal absorbs all the gains or losses from z shocks. In contrast, in the formulation in the main text, (3)–(5), it is the reverse: the agent's utility varies with z' and its wealth does not. Since agent wealth is a negative scalar multiple of the principal's utility (profits) this means that the principal's welfare is made independent of z'. Exactly the reverse as in (P3). To see this even more clearly, shut down residual productivity shocks, $\varepsilon = 1$ with probability one. Then the formulation in the main text, (3)–(5) is an income fluctuations problem, like Schechtman and Escudero (1977), Aiyagari (1994) or other Bewley models. But (P3) is just perfect insurance, with a risk neutral principal.

D Numerical Solution: Optimal Contract with Lotteries

When solving the optimal contract under moral hazard (3)–(5) numerically, we allow for lotteries as in Phelan and Townsend (1991). This section formulates the associated dynamic program.

D.1 Simplification

Capital, labor and occupational choice only enter the problem in (3) through the budget constraint (4). We can make use of this fact to reduce the number of choice variables in (3) from six $(e, x, k, \ell, c(\varepsilon), a'(\varepsilon))$ to three $(e, c(\varepsilon), a'(\varepsilon))$.

Entrepreneurs solve the following profit maximization problem.

$$\bar{\Pi}(z, e; w, r) = \max_{k,\ell} \ \bar{\varepsilon}(e) z f(k,\ell) - (r+\delta)k - w\ell, \quad \bar{\varepsilon}(e) \equiv \sum_{\varepsilon} p(\varepsilon|e)\varepsilon.$$

Note in particular that capital k and labor ℓ are chosen *before* residual productivity ε is realized (see the timeline in Figure 1). With the functional form assumption in (12), the first-order conditions are

$$\alpha z \bar{\varepsilon}(e) k^{\alpha-1}\ell^{\gamma} = r+\delta, \quad \gamma z \bar{\varepsilon}(e) k^{\alpha}\ell^{\gamma-1} = w$$

These can be solved for the optimal factor demands given effort, e, talent, z and factor prices w and r.

$$k^*(e, z; w, r) = (\bar{\varepsilon}(e)z)^{\frac{1}{1-\alpha-\gamma}} \left(\frac{\alpha}{r+\delta}\right)^{\frac{1-\gamma}{1-\alpha-\gamma}} \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\alpha-\gamma}}$$

$$\ell^*(e, z; w, r) = (\bar{\varepsilon}(e)z)^{\frac{1}{1-\alpha-\gamma}} \left(\frac{\alpha}{r+\delta}\right)^{\frac{\alpha}{1-\alpha-\gamma}} \left(\frac{\gamma}{w}\right)^{\frac{1-\alpha}{1-\alpha-\gamma}}$$

Realized (as opposed to expected) profits are

$$\Pi(\varepsilon, z, e; w, r) = z\varepsilon k(e, z; w, r)^{\alpha}\ell(e, z; w, r)^{\gamma} - w\ell(e, z; w, r) - (r+\delta)k(e, z; w, r)$$

Substituting back in from the factor demands, realized profits are

$$\Pi(\varepsilon, z, e; w, r) = \left(\frac{\varepsilon}{\bar{\varepsilon}(e)} - \alpha - \gamma\right)(z\bar{\varepsilon}(e))^{\frac{1}{1-\alpha-\gamma}}\left(\frac{\alpha}{r+\delta}\right)^{\frac{\alpha}{1-\alpha-\gamma}}\left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\alpha-\gamma}} \tag{22}$$

and expected profits are

$$\bar{\Pi}(z, e; w, r) = (1-\alpha-\gamma)\,(z\bar{\varepsilon}(e))^{\frac{1}{1-\alpha-\gamma}}\left(\frac{\alpha}{r+\delta}\right)^{\frac{\alpha}{1-\alpha-\gamma}}\left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\alpha-\gamma}} \tag{23}$$

The optimal occupational choice satisfies (note that agents choose an occupation *before* ε is realized):

$$x(z, e; w, r) = \arg\max_{x}\left\{x\bar{\Pi}(z, e; w, r) + (1-x)w\bar{\varepsilon}(e)\right\}$$

Given a realization of ε, those who choose to be entrepreneurs realize profits of (22) and those who choose to be workers realize a labor income of $w\varepsilon$. Therefore, realized (as opposed to expected) surplus is

$$S(\varepsilon, z, e; w, r) = x(z, e; w, r)\Pi(\varepsilon, z, e; w, r) + (1 - x(e, z; w, r))w\varepsilon.$$

Using these simplifications, the budget constraint (4) can then be written as

$$\sum_{\varepsilon} p(\varepsilon|e)\left\{c(\varepsilon) + a'(\varepsilon)\right\} = \sum_{\varepsilon} p(\varepsilon|e)S(\varepsilon, z, e; w, r) + (1+r)a. \tag{24}$$

As already noted, the advantage of this formulation is that it features three rather than six choice variables.

D.2 Linear Programming Representation

A contract between the intermediary and a household specifies a probability distribution over the vector

$$(c, \varepsilon, e, a')$$

given (a, z). Denote this probability distribution by $\pi(c, \varepsilon, e, a'|a, z)$. The associated dynamic program then is a linear programming problem where the choice variables are the probabilities $\pi(c, \varepsilon, e, a'|a, z)$:

$$v(a, z) = \max_{\pi(c,\varepsilon,e,a'|a,z)} \sum_{c,\varepsilon,e,a'} \pi(c, \varepsilon, e, a'|a, z)\left\{u(c, e) + \beta\mathbb{E}v(a', z')\right\} \quad \text{s.t.} \tag{25}$$

$$\sum_{c,\varepsilon,e,a'} \pi(c,\varepsilon,e,a'|a,z)\left\{a'+c\right\} = \sum_{c,\varepsilon,e,a'} \pi(c,\varepsilon,e,a'|a,z) S(\varepsilon,e,z;w,r) + (1+r)a. \tag{26}$$

$$\sum_{c,\varepsilon,a'} \pi(c,\varepsilon,e,a'|a,z)\left\{u(c,e)+\beta\mathbb{E}v(a',z')\right\} \geq \sum_{c,\varepsilon,a'} \pi(c,\varepsilon,e,a'|a,z)\frac{p(\varepsilon|\hat{e})}{p(\varepsilon|e)}\left\{u(c,\hat{e})+\beta\mathbb{E}v(a',z')\right\}\ \forall e,\hat{e}$$

$$\sum_{c,a'} \pi(c,\varepsilon,e,a'|a,z) = p(\varepsilon|e)\sum_{c,\varepsilon,a'} \pi(c,\varepsilon,e,a'|a,z), \quad \forall \varepsilon, e \tag{27}$$

(26) is the analogue of (24). The set of constraints (27) are the Bayes consistency constraints.[45]

D.3 Bounds on Consumption Grid

To solve the optimal contracting problem, we follow Prescott and Townsend (1984) and Phelan and Townsend (1991) and constrain all variables to lie on discrete grids. In order for the discretized dynamic programming problem to be a good approximation to our original problem, it turns out to be important to work with relatively fine grids, particularly for consumption. To achieve this with a limited number of grid points, we choose as tight an upper bound on the consumption grid as possible and adjust it when prices change. In particular, given (w,r), the upper bound is chosen as

$$\bar{c}(w,r) = r\bar{a} + \max\{\Pi(\varepsilon^H,\bar{z},\bar{e};w,r), w\varepsilon^H\},$$

for any given (w,r), where $\underline{a},\bar{a}$ and so on are the lower and upper bounds on the grids for wealth and other variables, and where the profit function Π is defined in (22). These are the minimum and maximum levels of consumption that can be sustained if the agent were to choose $a'(\varepsilon)=a$ in (3). Note that this bound is tighter than what is typically chosen in the literature. After solving the dynamic programming problem, we verify that consumption never hits the upper bound. Table 5 lists our choices of grids.

[45] (27) is derived from the timing of the problem as follows. A lottery with probabilities $\Pr(e)$ first determines an effort, e, for each household. Then a second lottery with probabilities $\Pr(c,\varepsilon,a'|e)$ determines the remaining variables. Of course, nature plays a role in this second lottery since the conditional probabilities $p(\varepsilon|e)$ are technologically determined. It is therefore required that

$$\sum_{c,a'} \Pr(c,\varepsilon,a'|e) = p(\varepsilon|e). \tag{28}$$

We have that

$$\Pr(c,\varepsilon,a'|e) = \frac{\pi(c,\varepsilon,e,a')}{\sum_{c,\varepsilon,a'}\pi(c,\varepsilon,e,a')} \tag{29}$$

Combining (28) and (29), we have

$$\frac{\sum_{c,a'}\pi(c,\varepsilon,e,a')}{\sum_{c,a',\varepsilon}\pi(c,\varepsilon,e,a')} = p(\varepsilon|e),$$

which is (27) above.

Table 5: Variable Grids

Variable	grid size	grid range
Wealth, a	30	[0, 200]
Ability, z	15	$[1, 4]$
Consumption, c	30	$[0.00001, \bar{c}(w, r)]$
Efficiency, ε	2	[0, 2]
Effort, e	2	[0.1, 1]

E More Details on Moral Hazard vs. Limited Commitment

This Appendix summarizes additional implications of moral hazard for individual choices and contrasts them with those of limited commitment. We relegated these to an Appendix because many of these, particularly for limited commitment, are already well understood from the existing literature.

E.1 Saving Behavior

We first present some analytic results that characterize differences in individual saving behavior in the two regimes. These are variants of well-known results in the literature.

Lemma 1 *Let $u(c, e) = U(c) - V(e)$. Solutions to the optimal contracting problem under moral hazard* (3)–(5)*, satisfy*

$$U'(c_{it}) = \beta(1 + r_{t+1})\mathbb{E}_{z,t}\left(\mathbb{E}_{\varepsilon,t}\frac{1}{U'(c_{it+1})}\right)^{-1} \tag{30}$$

where $\mathbb{E}_{z,t}$ and $\mathbb{E}_{\varepsilon,t}$ denote the time t expectation over future values of z and ε.

This is a variant of the inverse Euler equation derived in Rogerson (1985), Ligon (1998) and Golosov, Kocherlakota and Tsyvinski (2003) among others. With a degenerate distribution for ability, z, our equation collapses to the standard inverse Euler equation. The reason our equation differs from the latter is that we have assumed that ability, z, is not insurable in the sense that asset payoffs are not contingent on the realization of z (see footnote 13). Our equation is therefore a "hybrid" of an Euler equation in an incomplete markets setting and the inverse Euler equation under moral hazard.

If the incentive compatibility constraint (5) is binding, marginal utilities are not equalized across realizations of ε. One well known implication of (30) is that in this case[46]

$$U'(c_{it}) < \beta(1+r_{t+1})\mathbb{E}_{z,t}\mathbb{E}_{\varepsilon,t}U'(c_{it+1}). \tag{31}$$

The implication of this inequality is that when the incentive constraint binds, individuals are *saving constrained*. It is important to note that such saving constraints are a feature of the *optimal contract*.[47] The intuition is that under moral hazard there is an additional marginal cost of saving an extra dollar from period t to period $t+1$: in period $t+1$ an individual works less in response to any given compensation schedule. Therefore the optimal contract discourages savings whenever the incentive compatibility constraint (5) binds.

With limited commitment, the Euler equation is instead[48]

$$U'(c_{it}) = \beta\mathbb{E}_{z,t}\left[U'(c_{it+1})(1+r_{t+1}) + \nu_{it+1}\lambda\right]$$

where ν_{it+1} is the Lagrange multiplier on the collateral constraint (6). If this constraint binds, then

$$U'(c_{it}) > \beta(1+r_{t+1})\mathbb{E}_{z,t}U'(c_{it+1}). \tag{32}$$

Contrasting (31) for moral hazard and (32) for limited commitment, we can see that in the moral hazard regime individuals are *savings constrained* and in the limited commitment regime, they are instead *borrowing constrained*.[49] Finally, note that under limited commitment only the savings of entrepreneurs are distorted because only they face the collateral constraint (6). In contrast, under moral hazard the savings decision of both entrepreneurs and workers is distorted

[46]This follows because by Jensen's inequality ($1/U'(c_{it+1})$ is a convex function of $U'(c_{it+1})$)

$$\mathbb{E}_{\varepsilon,t}\frac{1}{U'(c_{it+1})} > \frac{1}{\mathbb{E}_{\varepsilon,t}U'(c_{it+1})}.$$

[47]Some readers may have had the opposite intuition, namely that moral hazard reduces insurance thereby strengthening precautionary motives for saving. But given that individuals' actions are governed by an optimal contract, the inverse the Euler equation says that this is not the case. See Rogerson (1985), Ligon (1998) and Golosov, Kocherlakota and Tsyvinski (2003) for more detailed discussions of this idea.

[48]Note that in contrast to (30) no expectation over ε is taken here. This is because there is perfect insurance on ε. Therefore marginal utilities are equalized across ε realizations. More formally, denote by $c(\varepsilon, z, a)$ consumption of an individual who has experienced shocks ε and z and has wealth a. Then $U'(c(\varepsilon, z, a)) = \psi(a, z)$ for all ε, where $\psi(a, z)$ is the Lagrange multiplier on the budget constraint in (4). Since this is true for all ε realizations, of course also $\mathbb{E}_\varepsilon U'(c(\varepsilon, z, a)) = \psi(a, z)$.

[49]In the case where the corresponding constraints do not bind, both (31) and (32) collapse to the standard Euler equation under incomplete markets

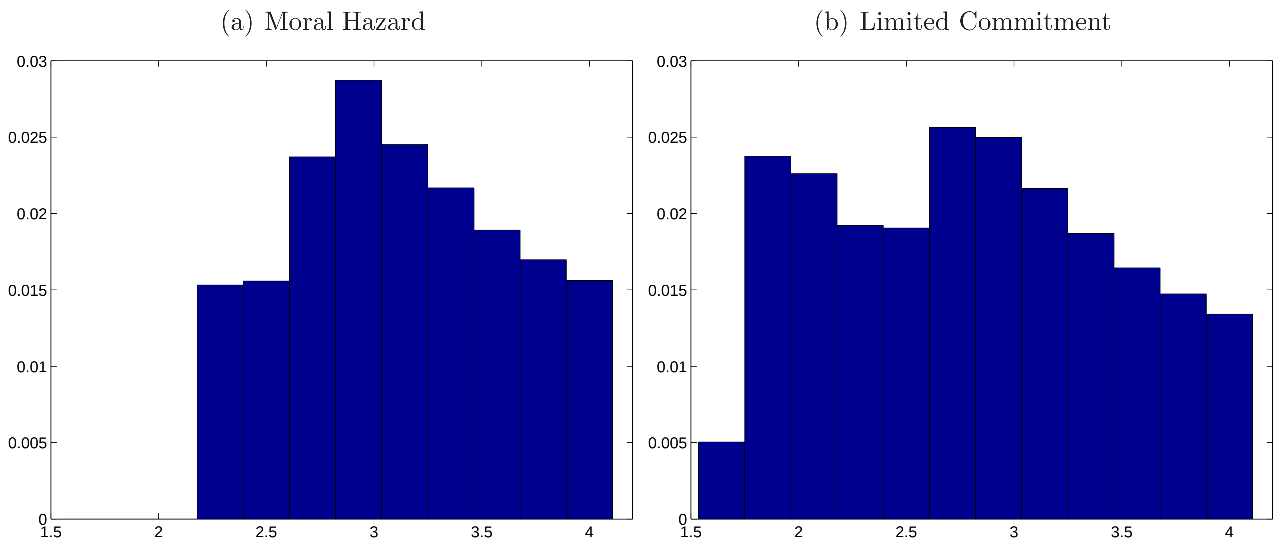
$$U'(c_{it}) = \beta(1+r_{t+1})\mathbb{E}_{z,t}U'(c_{it+1}).$$

because both face the incentive compatibility constraint (5). As discussed in the main text, this is reflected in the equilibrium interest rate (see Table 2). Individual savings behavior is one prediction in which the two regimes differ dramatically.

E.2 Cross-Section of Firms and Households

Apart from the differential effects of the two frictions on entrepreneurial effort, capital misallocation and occupational choose (Figures 4 to 5), there are also various other differences between the two regimes. The difference in occupational choice in the two economies, shown in Figure 5, immediately implies differences in the average entrepreneurial ability z of *active* entrepreneurs. Figure 10 displays the distributions of entrepreneurial abilities z of active entrepreneurs in the two economies. In the moral hazard economy, selection into entrepreneurship is more positive so that active entrepreneurs are more able on average. This is a force towards higher firm level TFP.

Figure 10: Distribution of Entrepreneurial Ability.

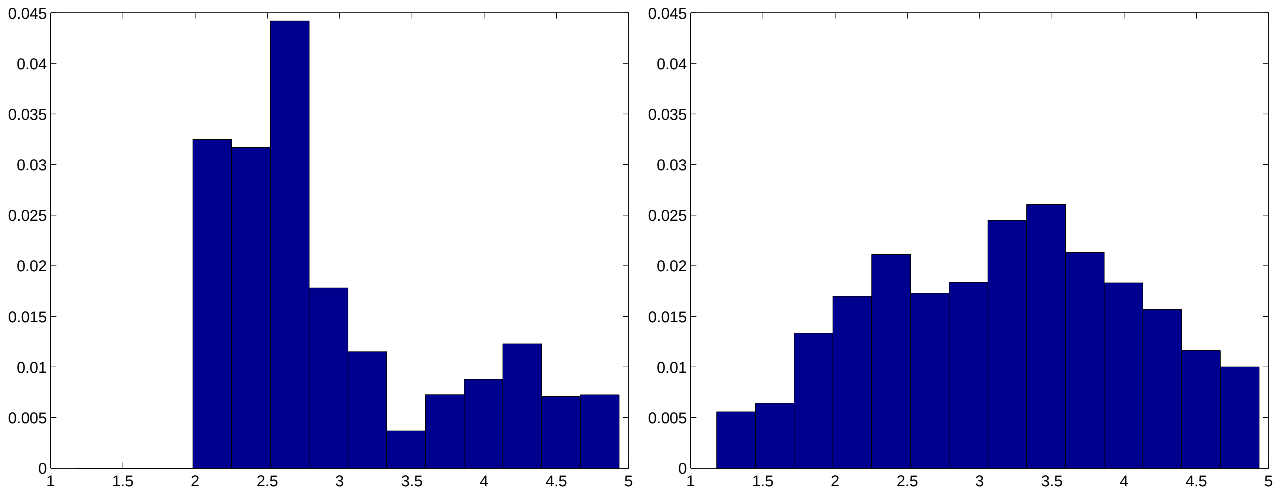


Figures 5 to 10 have shown that under moral hazard, entrepreneurs exert less effort but are more able on average. These two properties are jointly reflected in the distribution of observed firm-level TFP graphed in Figure 11. Recall that firm-level TFP is the product of "ability" and "residual productivity" and the latter depends on effort with probability distribution $p(\varepsilon|e)$. Ex-ante firm-level TFP is then given by $z\bar{\varepsilon}(e)$ where $\bar{\varepsilon}(e) \equiv \sum_{\varepsilon} \varepsilon p(\varepsilon|e)$ is expected residual productivity given an effort choice, e.[50] As shown in Figures 5 and 10, ability is higher in the

[50]To be clear, we here assume that capital and *hired* labor are accurately measured and accounted for as inputs into production, but that entrepreneurial effort is not. That is, each entrepreneur works full time at his firm and his time is counted as part of his firm's labor input. But effort is unobserved, implying that low

Figure 11: Distribution of Firm-level TFP.

(a) Moral Hazard (b) Limited Commitment



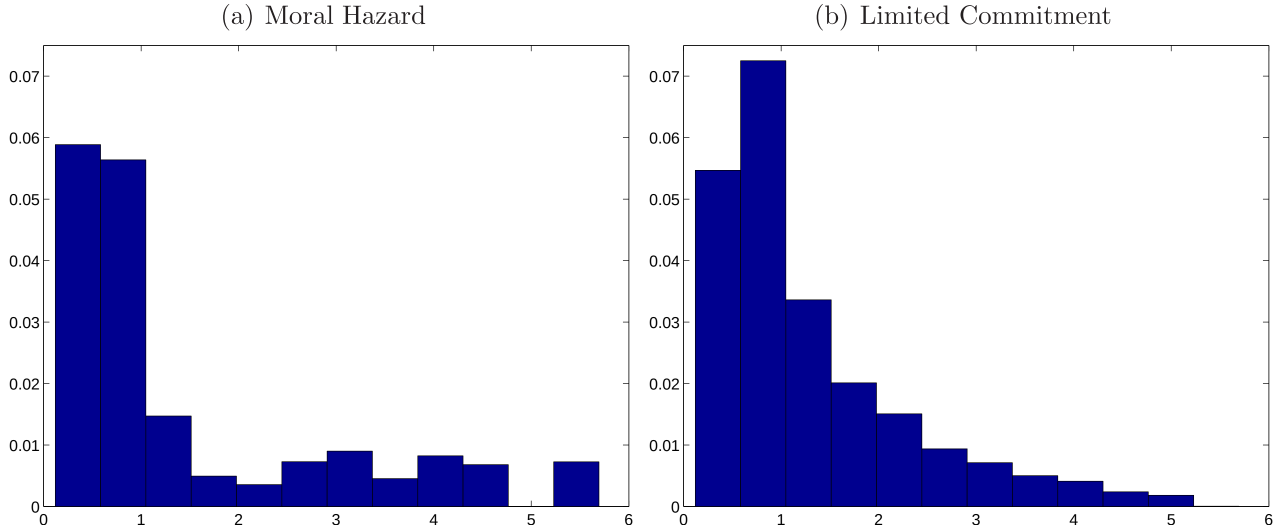
moral hazard economy. But as shown in Figure 10, effort is lower which results in lower realizations of residual productivity. As a result of these two offsetting forces, firm-level TFP is more dispersed in the limited commitment economy. Some high ability entrepreneurs additionally exert high effort and so have very high measured productivity; but there are also some low ability entrepreneurs (Figures 5 and 10) that additionally exert low effort and are hence very unproductive.

Consistent with these findings, Figure 12 plots the distribution of firm size as measured by a firm's number of employees. Firm size is similarly dispersed in both the moral hazard and limited commitment regimes. This is the outcome of two countervailing effects: on one hand, the productivity distribution is more dispersed under limited commitment than moral hazard which leads to more dispersion in firm size. On the other hand, a firm's capital stock is constrained by a collateral constraint under limited commitment, leading to less dispersion in limited commitment than moral hazard.

These results have important implications for measurement. For instance, consider an econometrician examining data generated by the moral hazard economy who measures gaps in marginal products of capital across individual firms. This econometrician would observe no capital misallocation and may therefore (erroneously) conclude that there is no friction in the capital market.

effort results in low measured firm-level TFP. Alternatively, we could have assumed that a similar measurement problem also applies to *hired* labor. In this case, losses from moral hazard in measured firm-level TFP would be further amplified.

Figure 12: Firm-Size (Employee) Distribution.



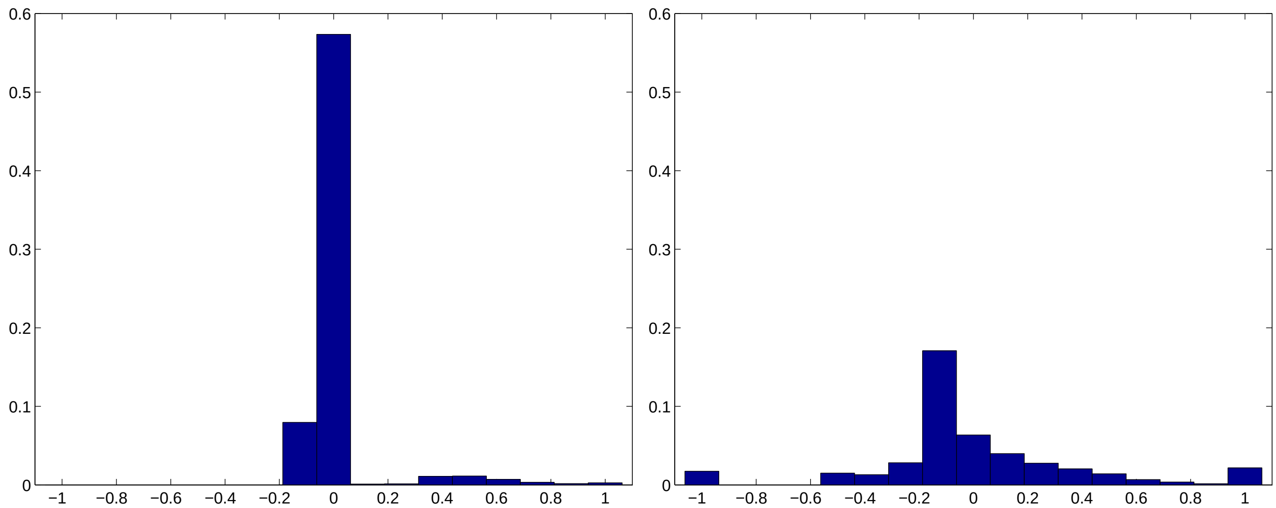
E.3 Dynamics of Firms and Households

Individual savings behavior in the two economies also differs, in particular the speed of individual transitions. One convenient way of summarizing this speed of transition is to compare the eigenvalues of the transition matrix $\Pr(a', z'|a, z)$ defined in (7) for the two economies. The eigenvalue governing the speed of convergence in the limited commitment economy is 0.9396 with a corresponding half life of $-\log(2)/\log(0.9396) \approx 11.1$ years whereas in the moral hazard economy this eigenvalue is 0.9823 which implies a half-life of 38.8 years.[51] Individual transitions are therefore much slower in the moral hazard economy. This is also shown in Figure 13 which plots the distribution of wealth growth rates in the two economies. In the moral hazard economy, most wealth growth rates are close to zero. In contrast, with limited commitment wealth growth rates are much more dispersed. This is a direct consequence of the difference in the dispersion of marginal products of capital in the two regimes (Figure 4). A high marginal product of capital due to a binding borrowing constraint also implies a high return to wealth accumulation (the Lagrange multiplier in the Euler equation) and therefore leads to faster wealth accumulation. High dispersion in marginal products therefore implies high dispersion in wealth growth rates. The finding that individual transitions are fast under limited commitment is related to the "conditional monotonicity" result of Albuquerque and Hopenhayn (2004) that at any point in time wealth and input levels are not lower than in any past state of nature characterized by the same productivity level. In contrast, under moral hazard low realizations of residual productivity lead in general to a wealth decline and individual wealth or promised utility moves slowly as output-dependent penalties and awards are spread into the future (see

[51]The speed of convergence is determined by the largest eigenvalue that is less than one (see e.g. Stokey, Lucas and Prescott, 1989).

Figure 13: Distribution of Wealth Growth Rates

(a) Moral Hazard (b) Limited Commitment



for example Phelan and Townsend, 1991; Karaivanov and Townsend, 2014).

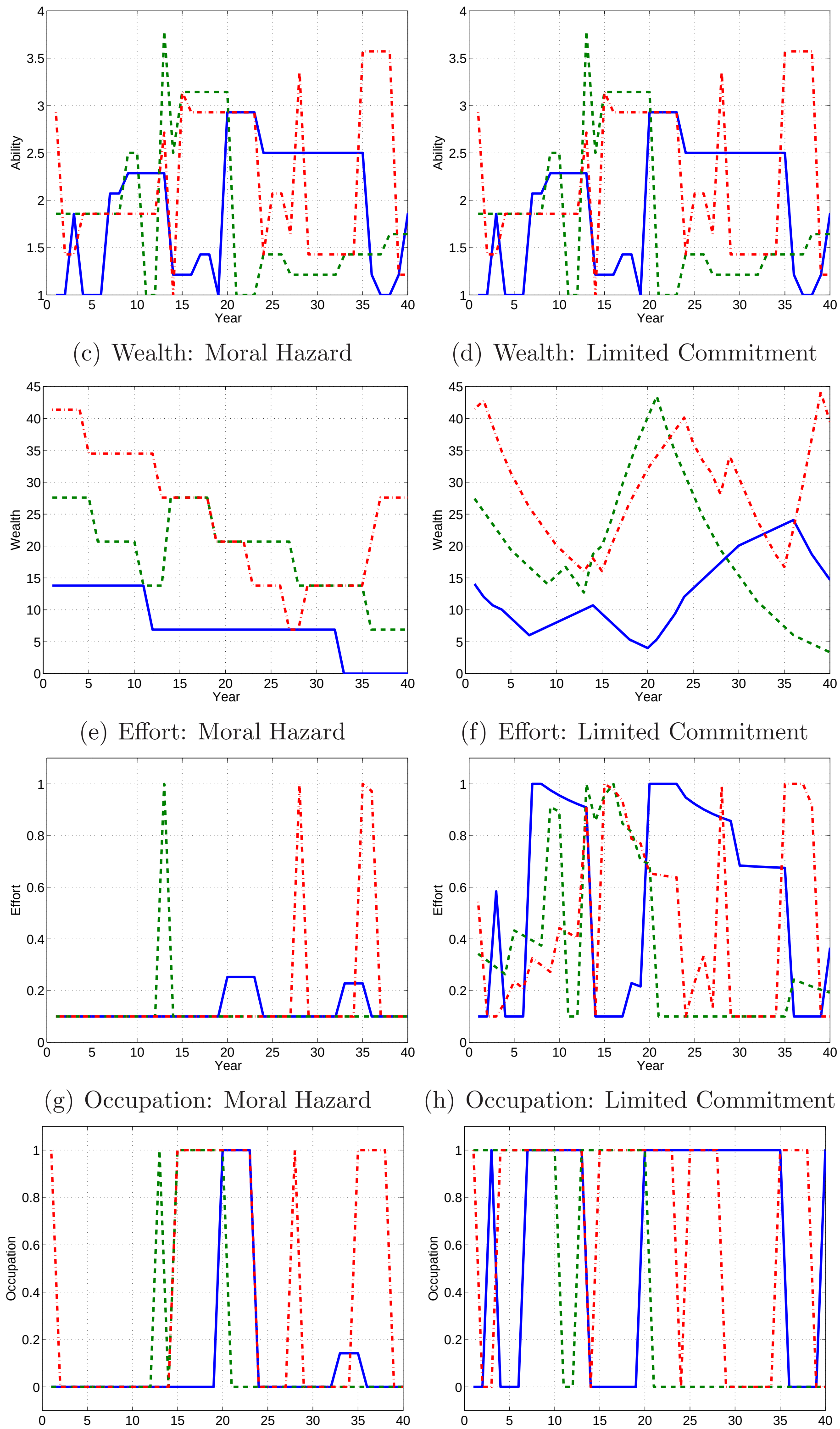
To further contrast individual behavior in the two regimes, Figure 14 displays sample paths of ability, wealth, effort and occupational choice for three typical individuals. The individuals start with the same level of wealth and experience the same sequence of productivity shocks. The only difference is the financial regime they operate in. Panels (c) and (d) show that individuals in the limited commitment regime accumulate and decumulate wealth at a faster speed than those in the moral hazard regime (consistent with Figure 13). Panels (e) and (f) show that individuals in the moral hazard regime exert lower effort, again as expected. Finally, panels (g) and (h) show that there is more occupational mobility in the limited commitment regime. This is because under limited commitment wealth is a more important determinant of entrepreneurship (see Figure 5), and wealth moves around more (panels (c) and (d)).

Wealth and income inequality also differ in the two regimes. Figure 15 reports the wealth Lorenz curves for the two regimes. It can be seen that wealth inequality in higher in the limited commitment regime. This is a direct consequence of the bigger dispersion in marginal products of capital in Figure 4, which leads to high wealth concentration in the hands of high productivity entrepreneurs trying to save themselves out of borrowing constraints. Similarly, Figure 16 plots the income Lorenz curves. Again, inequality is higher under limited commitment than under moral hazard.

References

Abraham, Arpad, and Nicola Pavoni. 2005. "The Efficient Allocation of Consumption

Figure 14: Typical Histories



(a) Ability: Moral Hazard (b) Ability: Limited Commitment

(c) Wealth: Moral Hazard (d) Wealth: Limited Commitment

(e) Effort: Moral Hazard (f) Effort: Limited Commitment

(g) Occupation: Moral Hazard (h) Occupation: Limited Commitment

Figure 15: Wealth Lorenz Curves

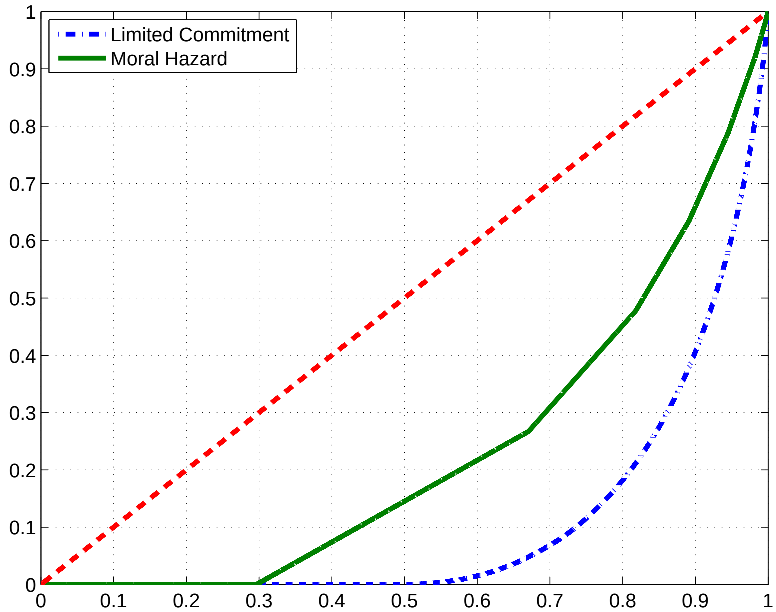
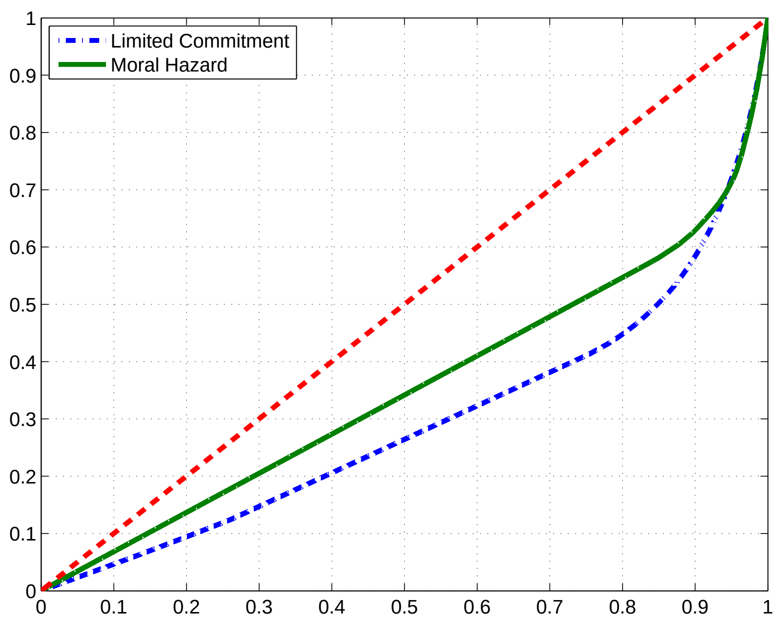


Figure 16: Income Lorenz Curves



under Moral Hazard and Hidden Access to the Credit Market." *Journal of the European Economic Association*, 3(2-3): 370–381.

Aghion, Philippe, and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *The Review of Economic Studies*, 64: 151–172.

Ahlin, Christian, and Robert M. Townsend. 2007. "Using Repayment Data to Test Across Models of Joint Liability Lending." *Economic Journal*, 117(517): F11–F51.

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

Albuquerque, Rui, and Hugo Hopenhayn. 2004. "Optimal Lending Contracts and Firm Dynamics." *Review of Economic Studies*, 71: 285315.

Attanasio, Orazio P., and Nicola Pavoni. 2011. "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption." *Econometrica*, 79(4): 1027–1068.

Banerjee, Abhijit V., and Benjamin Moll. 2010. "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2(1): 189–206.

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth.* , ed. Philippe Aghion and Steven Durlauf.

Banerjee, Abhijit V, Timothy Besley, and Timothy W Guinnane. 1994. "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test." *The Quarterly Journal of Economics*, 109(2): 491–515.

Beck, Thorsten, A. Demirguc-Kunt, and Ross Levine. 2000. "A New Database on the Structure and Development of the Financial Sector." *World Bank Economic Review*, 14: 597605.

Besley, Timothy, and Stephen Coate. 1995. "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics*, 46(1): 1–18.

Broer, Tobias. 2013. "The Wrong Shape of Insurance? What Cross-Sectional Distributions Tell Us about Models of Consumption Smoothing." *American Economic Journal: Macroeconomics*, 5(4): 107–40.

Buera, Francisco J., and Yongseok Shin. 2011. "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks." *Journal of Economic Theory*, 146(3): 845–862.

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

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

Caselli, Francesco, and Nicola Gennaioli. 2013. "Dynastic Management." *Economic Inquiry*, 51(1): 971–996.

Castro, Rui, Gian Luca Clementi, and Glenn Macdonald. 2009. "Legal Institutions, Sectoral Heterogeneity, and Economic Development." *Review of Economic Studies*, 76(2): 529–561.

Clementi, Gian Luca, and Hugo A. Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics." *Quarterly Journal of Economics*, 229–265.

Cole, Harold L., Jeremy Greenwood, and Juan M. Sanchez. 2012. "Why Doesn't Technology Flow from Rich to Poor Countries?" University of Pennsylvania mimeo.

Collard-Wexler, Allan, John Asker, and Jan DeLoecker. 2011. "Productivity Volatility and the Misallocation of Resources in Developing Economies." National Bureau of Economic Research, Inc NBER Working Papers 17175.

Doepke, Matthias, and Robert M. Townsend. 2006. "Dynamic mechanism design with hidden income and hidden actions." *Journal of Economic Theory*, 126(1): 235–285.

Evans, David, and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy*, 97(4): 808–827.

Ghatak, Maitreesh. 1999. "Group lending, local information and peer selection." *Journal of Development Economics*, 60(1): 27–50.

Ghatak, Maitreesh, Massimo Morelli, and Tomas Sjostrom. 2001. "Occupational Choice and Dynamic Incentives." *Review of Economic Studies*, 68(4): 781–810.

Ghosh, Parikshit, Dilip Mookherjee, and Debraj Ray. 2000. "Credit Rationing in Developing Countries: An Overview of the Theory." In *A Reader in Development Economics.* , ed. Dilip Mookherjee and Debraj Ray. Blackwell.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569–587.

Gourio, Francois. 2008. "Estimating Firm-Level Risk." Boston University Working Paper.

Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2010*a*. “Financing Development: The Role of Information Costs.” *American Economic Review*, 100: 18751891.

Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2010*b*. “Quantifying the Impact of Financial Development on Economic Development.” Economie d’Avant Garde Economie d’Avant Garde Research Reports 17.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman. 1988. “Investment, Capacity Utilization, and the Real Business Cycle.” *American Economic Review*, 78(3): 402–17.

Guerrieri, Veronica, and Guido Lorenzoni. 2011. “Credit Crises, Precautionary Savings and the Liquidity Trap.” University of Chicago mimeo.

Harris, John R, and Michael P Todaro. 1970. “Migration, Unemployment & Development: A Two-Sector Analysis.” *American Economic Review*, 60(1): 126–42.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S Rosen. 1994. “Sticking It Out: Entrepreneurial Survival and Liquidity Constraints.” *Journal of Political Economy*, 102(1): 53–75.

Jeong, Hyeok, and Robert Townsend. 2007. “Sources of TFP Growth: Occupational Choice and Financial Deepening.” *Economic Theory*, 32: 179–221.

Karaivanov, Alexander, and Robert Townsend. 2014. “Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes.” *Econometrica*. forthcoming.

Kareken, John H, and Neil Wallace. 1978. “Deposit Insurance and Bank Regulation: A Partial-Equilibrium Exposition.” *The Journal of Business*, 51(3): 413–38.

Kinnan, Cynthia. 2012. “Distinguishing Barriers to Insurance in Thai Villages.” Northwestern University Working Paper.

Kinnan, Cynthia, and Robert Townsend. 2012. “Kinship and Financial Networks, Formal Financial Access, and Risk Reduction.” *American Economic Review*, 102(3): 289–93.

Krueger, Dirk, and Fabrizio Perri. 2011. “How do Households Respond to Income Shocks?” Working Paper.

Lewis, William Arthur. 1954. “Economic Development with Unlimited Supplies of Labor.” In *Manchester School of Economics and Social Studies*.

Ligon, Ethan. 1998. "Risk Sharing and Information in Village Economics." *Review of Economic Studies*, 65(4): 847–64.

Lucas, Robert E. 2003. "Macroeconomic Priorities." *American Economic Review*, 93(1): 1–14.

Martin, Alberto, and Filippo Taddei. 2012. "International Capital Flows and Credit Market Imperfections: a Tale of Two Frictions." Barcelona Graduate School of Economics Working Paper.

Maurer, Klaus, and Hans Dieter Seibel. 2001. "Agricultural Development Bank Reform: The Case of Unit Banking System of Bank Rakyat Indonesia (BRI)." International Fund for Agricultural Development Working Paper No. B5 (Doc. #48881).

Meisenzahl, Ralf R. 2011. "Verifying the state of financing constraints: evidence from U.S. business credit contracts." Board of Governors of the Federal Reserve System (U.S.) Finance and Economics Discussion Series 2011-04.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review*. forthcoming.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?" *American Economic Review*. forthcoming.

Morten, Melanie. 2013. "Temporary Migration and Endogenous Risk Sharing in Village India." Working Paper.

Paulson, Anna L., Robert M. Townsend, and Alexander Karaivanov. 2006. "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship." *Journal of Political Economy*, 114(1): 100–144.

Phelan, Christopher, and Robert M Townsend. 1991. "Computing Multi-period, Information-Constrained Optima." *Review of Economic Studies*, 58(5): 853–81.

Piketty, Thomas. 1997. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *The Review of Economic Studies*, 64: 173–189.

Prescott, Edward C. 2004. "Why do Americans work so much more than Europeans?" *Quarterly Review*, , (Jul): 2–13.

Prescott, Edward C, and Robert M Townsend. 1984. "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard." *Econometrica*, 52(1): 21–45.

Rampini, Adriano A. 2004. "Entrepreneurial activity, risk, and the business cycle." *Journal of Monetary Economics*, 51(3): 555–573.

Rogerson, William P. 1985. "Repeated Moral Hazard." *Econometrica*, 53(1): 69–76.

Samphantharak, Krislert, and Robert M. Townsend. 2009. "Households as Corporate Firms: Constructing Financial Statements from Integrated Household Surveys." Econometric Society Monograph Series.

Schechtman, Jack, and Vera L. S. Escudero. 1977. "Some results on 'an income fluctuation problem'." *Journal of Economic Theory*, 16(2): 151–166.

Schmid, Lukas. 2012. "A Quantitative Dynamic Agency Model of Financing Constraints." Working Paper.

Shimer, Robert. 2010. *Labor Markets and Business Cycles.* Princeton University Press.

Shourideh, Ali. 2012. "Optimal Taxation of Capital Income: A Mirrleesian Approach to Capital Accumulation." Wharton Working Paper.

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies*, 54(4): 599–617.

Stiglitz, Joseph E. 1990. "Peer Monitoring and Credit Markets." *World Bank Economic Review*, 4(3): 351–66.

Stokey, Nancy L., Robert E. Jr. Lucas, and Edward C. Prescott. 1989. *Recursive Methods in Economic Dynamics.* Harvard University Press.

Townsend, Robert M. 1979. "Optimal contracts and competitive markets with costly state verification." *Journal of Economic Theory*, 21(2): 265–293.