Asset Pricing with Endogenously Uninsurable Tail Risk*

Hengjie Ai and Anmol Bhandari

November 29, 2020

This paper studies asset pricing and labor market dynamics when idiosyncratic risk to human capital is not fully insurable. Firms use long-term contracts to provide insurance to workers, but neither side can fully commit; furthermore, owing to costly and unobservable retention effort, worker-firm relationships have endogenous durations. Uninsured tail risk in labor earnings arises as a part of an optimal risk-sharing scheme. In equilibrium, exposure to the tail risk generates higher aggregate risk premia and higher return volatility. Consistent with data, firm-level labor share predicts both future returns and pass-throughs of firm-level shocks to labor compensation.

Key words: Equity premium puzzle, dynamic contracting, tail risk, limited commitment

*Hengjie Ai (hengjie@umn.edu) is affiliated with the Carlson School of Management, University of Minnesota; Anmol Bhandari (bhandari@umn.edu) is at the Department of Economics, University of Minnesota. The authors would like to thank V.V. Chari, YiLi Chien, David Evans, Barney Hatman-Glaser, Urban Jermann, Hanno Lustig, Tom Sargent, Lukas Schmid, Stijn Van Nieuwerburgh, Jessica Wachter, and seminar participants at the 2016 NBER AP meetings, the Minnesota Macro Workshop, the 2017 Minnesota Macro Asset Pricing Conference, the 2017 SITE Workshop, Stanford University, the UBC winter finance conference, Booth Finance Workshop at the University of Chicago, the Money and Banking Workshop at University of Chicago, the University of North Carolina, the University of Rochester, the 2017 WFA Annual Meetings, the Wharton International Finance Workshop, and the Utah Winter Finance Conference for their helpful comments. The authors would also like to thank Jincheng Tong, Yuki Yao, and Chao Ying for their assistance and comments on the paper. We thank Joan Gieseke and James Holt for excellent editorial assistance.

1 Introduction

A key challenge for macro asset pricing theories is to account for the large magnitude of equity premia and their substantial variations over time and across firms. In this paper, we provide an asset pricing model with imperfect risk sharing to address these patterns in risk premia. Uninsured *tail* or downside risk in labor earnings arises as an outcome of optimal risk-sharing arrangements in presence of limited commitment and moral hazard. Time variation in that tail risk drives aggregate equity premium and the volatility of stock market returns. The model is also consistent with firm-level labor share predicting both future returns and pass-throughs of firm-level shocks to labor compensation. Overall, the paper provides a unified view of labor market risk and asset prices within a general equilibrium optimal contracting framework.

The setup consists of two types of agents: firm owners and workers. Firm owners are well diversified and use long-term compensation contracts to provide insurance to workers who face idiosyncratic shocks to their human capital. Two agency frictions distinguish our paper from a standard representative agent asset pricing model. First, neither firm owners nor workers can commit to contracts that yield continuation values lower than their outside options. Second, owing to costly and unobservable retention effort, worker-firm relationships have endogenous durations. We embed these contracting frictions in a general equilibrium setting with aggregate shocks to study their labor market and asset pricing implications.

Downside risk in labor earnings, a key feature of the data, is driven by firm-side limited commitment and further amplified by the presence of moral hazard. Compensation contracts providing perfect risk sharing would insure workers against idiosyncratic labor productivity shocks. But when firms cannot commit to matches that yield negative net present value, large drops in labor productivity are accompanied by reductions in worker earnings. In addition, because retention effort is not observable, firms have a lower incentive to keep workers when the firm-worker match is less profitable. Thus, separation risk is elevated after adverse productivity shocks.

In the general equilibrium, exposure to downside risk drives several of our asset pricing results. First, it generates a stochastic discount factor that is more volatile than that in an otherwise identical economy without agency frictions. A necessary condition is recursive utility with preference for early resolution of uncertainty and persistent and countercyclical idiosyncratic risk to worker human capital. During recessions, the anticipation of lack of risk sharing in the future raises workers' current marginal utilities. The optimal risk-sharing scheme compensates by allocating a higher share of aggregate output from firm owners to workers. Therefore, the labor share moves inversely with aggregate output. The countercyclicality of labor share translates into a procyclical consumption share of all unconstrained investors, including firm owners. This amplifies risk prices. In our quantitative analysis, owing to agency frictions, we find that Sharpe ratios are more than doubled.

The high volatility of the stochastic discount factor is further amplified by the moral hazard in firms' choice of retention effort. Higher expected returns during recessions lower valuation and weaken firms' incentive to retain workers, resulting in countercyclical separations. This feature of our model supplements a large macro labor literature that argues that discount rate variations are central in driving unemployment fluctuations. In our model, higher separations exacerbate tail risk and therefore the need for firm owners to provide insurance against aggregate shocks. This further raises equilibrium discount rates and amplifies risk prices.

Second, without relying on heteroskedastic aggregate shocks, our model produces substantial predictable variations in the risk premium, especially over long horizons. The dynamics of the pricing kernel depend on the fraction of firms that are likely to hit their limited commitment constraint. This introduces persistent variations in the volatility of the stochastic discount factor and makes returns predictable. In our model, regressions of returns for a claim to aggregate consumption on price-dividend ratios gives R-squares that are significant and increasing in horizon. Time variation in discount rates also amplifies the response of asset prices to aggregate shocks and further elevates the market equity premium.

Lastly, the above economic mechanism results in a significant heterogeneity in the cross section of expected equity returns and pass-through of firm-level shocks to labor compensation. Under the optimal contract, payments to workers insures them against aggregate productivity shocks making the residual capital income procyclical and thus more exposed to aggregate shocks. This delivers a form of operating leverage at the firm level. Firms that have experienced adverse idiosyncratic shocks have a higher fraction of their value promised to workers and are therefore more sensitive to aggregate shocks. As a result, they have lower valuation ratios and higher expected returns. Furthermore, firms with larger obligations to workers are more likely to hit the firm-side limited commitment constraint and are more likely to lower wage payments in response to an adverse idiosyncratic shock. We test these implications using CRSP/Compustat panel data and show that firm-level measures of labor share predict both future returns and pass-throughs of firm-level shocks to labor compensation.

Related literature This paper builds on the literature on limited commitment. Kehoe and Levine (1993) and Alvarez and Jermann (2000) develop a theory of incomplete markets based on one-sided limited commitment. Kocherlakota (1996) examines the implications of limited commitment on consumption risk sharing. Alvarez and Jermann (2001) and Chien and Lustig (2010) study the asset pricing implications of such environments. Rampini and

Viswanathan (2010, 2013) develop a theory of capital structure based on limited commitment. Rampini and Viswanathan (2017) study the trade-off between intertemporal financing needs and insurance across states in a household insurance problem with limited commitment.

Most of the above theory builds on the Kehoe and Levine (1993) framework and implies that agents who experience large positive income shocks have an incentive to default because they have better outside options. As a result, positive income shocks cannot be insured, while downside risk in income is perfectly insured. To be consistent with data on labor earnings which show that downside earning risks are uninsured, we model two-sided lack of commitment as in Thomas and Worrall (1988) and augment it with moral hazard. The firm-side limited commitment problem in our model has a similar structure to those studied in Bolton et al. (2019) and Ai and Li (2015) and Ai et al. (forthcoming). In addition, we add aggregate shocks and focus on the general equilibrium effects of the firm-side limited commitment and moral hazard that have not been studied before.[1]

Our paper is related to asset pricing models with incomplete markets. Krueger and Lustig (2010) provide theoretical conditions under which the presence of idiosyncratic risk is irrelevant for the market price of aggregate risks. Mankiw (1986) and Constantinides and Duffie (1996) demonstrate how countercyclical volatility in incomes amplifies aggregate risk premia in the general equilibrium. Schmidt (2015) and Constantinides and Ghosh (2014) calibrate such incomplete markets models to recent administrative data on earnings and show that higher moments of labor income shocks require a significant risk compensation. For tractability, the Constantinides and Duffie (1996) framework requires an assumption of independently distributed shocks to income growth, which rules out any insurance or trading of financial assets in equilibrium and imposes that exposure to aggregate shocks is same for all households.[2] In contrast to the above papers, we take an optimal contracting approach to micro-found incomplete markets and use empirical evidence on labor earnings dynamics to restrict the choice of the parameters governing agency frictions. Our model allows investors to access a rich set of state-contingent payoffs. We explicitly characterize history dependence in labor earnings risk and exposure to aggregate shocks.

Our paper is also related to the literature on asset pricing with heterogeneous preferences, for instance, Guvenen (2009), Garleanu et al. (2012), Garleanu and Panageas (2015), Veronesi (2019), Gomez (2019), and Santos and Veronesi (2020). See Panageas (2019) for a recent review of this literature. These papers share a feature that effective risk aversion is a wealth-

[1]Recently, several papers such as Tsuyuhara (2016), Abraham et al. (2017), and Lamadon (2016) study versions of long-term wage contracts with moral hazard. Lamadon (2016) allows for richer features such as worker and firm complementarities, and on-the-job search. However, none of these papers allow for aggregate risks or study asset pricing.

[2]Heaton and Lucas (1996) and Storesletten et al. (2007) are among the few papers that depart from the no-trade equilibrium to study risk premia in quantitative models with exogenously incomplete markets.

weighted average of individual risk aversions and generate time-varying equity premia when wealth shares fluctuate with aggregate shocks. In order to generate a sufficiently volatile stochastic discount factor, these settings still need to rely on high levels and large differences in risk aversion along with substantial movements in the wealth distribution. In contrast, our model uses homogeneous preferences with low risk aversion. Our model generates substantial movements in the volatility of stochastic discount factor because agency frictions amplify the risk exposure of the marginal agents who provide efficient insurance to rest of the economy.

Theoretical predictions of our model are also consistent with a recent literature that emphasizes the importance of labor-share dynamics in understanding asset prices. Our operating leverage results connect to insights in Danthine and Donaldson (2002) and Berk and Walden (2013). More recently, Favilukis and Lin (2016b) use models with sticky wages to demonstrate how countercyclical movements in labor shares help explain equity and credit risk premia in production economies. Our model's implication that variations in labor shares can account for a large fraction of aggregate stock market variations is consistent with the evidence in Greenwald et al. (2016, 2020) and Lettau et al. (2014).

Our computational method builds on Krusell and Smith (1998). Using techniques contributed by the dynamic contracting literature, such as Atkeson and Lucas (1992), we represent equilibrium allocations recursively by using a distribution of promised values as a state variable. However, in contrast to those papers, our environment has aggregate shocks, and the distribution of promised values responds to such shocks, even in an ergodic steady state.

The paper is organized as follows. In section 2, we describe the physical as well as the contracting environment, and formalize a recursive competitive equilibrium with long-term contracts. In section 3, we discuss the optimal contract. In section 4, we derive the asset pricing implications that arise from agency frictions and general equilibrium considerations. Finally, in sections 5 and 6, we present quantitative implications after calibrating to several aggregate and cross-sectional facts. Section 7 concludes.

2 Model

We start with the physical and contracting environment.

2.1 Setup

Demographics and endowments We consider a discrete time economy with $t = 0, 1, \ldots$. There are two groups of agents: a unit measure of firm owners and a unit measure of workers. Members of both groups have Epstein-Zin preferences with a common risk aversion γ and a common intertemporal elasticity of substitution (IES) ψ. In each period, workers die with probability $1-\kappa$, and a similar measure of new workers are born. This specification guarantees that the total measure of workers equals one at all times. Upon birth, a worker is endowed with one unit of human capital. Firm owners are endowed with a diversified portfolio of equity shares in all firms.

Production and human capital Production is organized within N firms. We assume that N is large so that firms are perfectly competitive. In any period, a worker is either unemployed or matched to a firm. A worker produces output only in the firm in which he is employed. If employed in period t, worker i with human capital $h_{i,t}$ produces output

$$y_{i,t} = Y_t h_{i,t},$$

where Y_t is the aggregate productivity. We assume $Y_0 = 1$, and for $t \geq 1$,

$$\ln Y_{t+1} = \ln Y_t + g_t,$$

where g_t is a finite state Markov process with a one-step transition matrix $\{\pi\left(g'|g\right)\}_{g,g'}$.

The evolution of worker human capital in the next period depends on whether the worker is employed or unemployed. The law of motion for the human capital of worker i who remains employed with firm j in period $t+1$ is

$$h_{i,t+1} = h_{i,t} e^{\eta_{j,t+1}+\varepsilon_{i,t+1}}, \tag{1}$$

where conditioning on the aggregate Markov state g_t, the firm component $\eta_{j,t}$ is i.i.d. across firms but common to all workers in a firm; the worker-specific shock $\varepsilon_{i,t}$ is i.i.d. across workers; and $\eta_{j,t}$ and $\varepsilon_{i,t}$ are mutually independent. We use $f\left(\eta, \varepsilon|g\right)$ for the conditional density of (η, ε) and normalize so that $\mathbb{E}\left[e^{\varepsilon_{i,t}}|g_t\right] = 1$ and $\mathbb{E}\left[e^{\eta_{j,t}}|g_t\right] = 1$. We use $z_{i,j,t} = (\eta_{j,t}, \varepsilon_{i,t})$ for match-specific shocks for the worker-firm pair (i, j) at time t. The human capital of a worker not matched with a firm—that is, a worker who becomes (or remains) unemployed in period $t+1$—depreciates deterministically according to

$$h_{i,t+1} = \lambda h_{i,t}, \tag{2}$$

where the parameter $\lambda < 1$ describes human capital obsolescence. In each period, unemployed workers receive unemployment benefit $bY_t h_{i,t}$, where b is a constant.

Matching and separation A match between a worker and a firm can end in two ways: stochastically upon the arrival of a separation shock, or voluntarily by the firm or the worker. Firms can influence the probability of separation by exerting costly effort. We interpret such effort as a proxy for investments in organization capital that allow firms to retain workers and help them accumulate human capital on the job. We denote the effort for keeping a worker at time t by θ_t and assume that the cost of effort per unit of output is specified by a function $A(\theta)$ with the first three derivatives strictly positive all $\theta \in (0,1)$. It is without loss of generality to denominate effort in probability units because the cost of effort is captured by the functional form of $A(\theta)$. In period $t+1$, conditioning on the survival of the match, both firms and workers can unilaterally initiate a separation. We denote such a voluntary separation decision by an indicator function δ_{t+1}, with $\delta_{t+1} = 0$ for separation.

Upon separation, a worker enters into unemployment. In each period, an unemployed worker receives an employment opportunity with probability $\chi \in (0,1)$. An employment opportunity enables a worker to access a labor market where firms offer long-term contracts. In addition to unemployed workers, newborn workers also have an employment opportunity. A worker with an employment opportunity can choose to establish a match with the firm that offers the most favorable contract. We assume that there is no cost for posting vacancies and all firms can compete for new workers.

Contracts Let τ denote the beginning of an employment relationship between a firm and a worker. A employment contract offered by a firm to a newly employed worker at time τ specifies: (i) net transfers or compensation from the firm to the worker $\{C_t\}_{t=\tau}^{\infty}$, (ii) firm's effort for keeping the match $\{\theta_t\}_{t=\tau}^{\infty}$, and (iii) match termination decisions $\{\delta_t\}_{t=\tau}^{\infty}$ for the duration of the match. Formally, an employment contract offered in period τ by firm j to worker i with human capital $h_{i,\tau}$ specifies $\{C_t, \theta_t, \delta_t\}_{t=\tau}^{\infty}$ as functions of aggregate and match-specific histories:

$$\mathscr{C}_{i,j,\tau} \equiv \left\{C_{i,j,t}\left(h_{i,\tau}, z_{i,j}^{\tau\to t}, g^t\right), \theta_{i,j,t}\left(h_{i,\tau}, z_{i,j}^{\tau\to t}, g^t\right), \delta_{i,j,t}\left(h_{i,\tau}, z_{i,j}^{\tau\to t}, g^t\right)\right\}_{t=\tau}^{\infty}. \tag{3}$$

We use the convention that superscript t denotes the history of shocks up to time t: $g^t = \{g_1, g_2, \cdots, g_t\}$, and superscript $\tau \to t$ denotes the history of shocks from time τ to t: $z_{i,j}^{\tau\to t} = \{z_{i,j,\tau+1}, z_{i,j,\tau+2}, \ldots, z_{i,j,t}\}$, with $z_{i,j}^{\tau\to\tau} = \emptyset$ representing the trivial history.

We have made two simplifying assumptions in the choice of the contract space. First, we restrict our attention to employment contracts and do not allow payments from firms to workers who are not matched with the firm. In appendix A, we show that because of two-sided limited commitment—an agency friction that we introduce below—firms are unable to insure unemployed workers or workers employed by other firms, even if they are allowed to offer insurance contracts to these workers.

Second, we assume that the terms of a new employment contract depend on worker-specific history only through the human capital at the time of employment. This is true as long as workers with an employment opportunity have no pre-existing obligations or claims and can extract full surplus from a new match. In our setup, limited commitment means that any pre-existing contracts can be costlessly reneged, while perfect competition among firms implies that the best contract indeed gives workers all the surplus. Therefore, it is without loss of generality to index new contracts with the human capital of the worker at the time when the match is formed. To simplify notion and exposition, we impose both of the above restrictions at the outset.

Firm value Let $\left\{\Lambda_t\left(g^t\right)\right\}_t$ denote the stochastic process for state prices; that is, $\Lambda_t\left(g^t\right)$ is the price at history g^t of one unit of consumption goods measured in period-0 consumption numeraire. For $t \geq \tau$, let $V_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}_{i,j,\tau}\right)$ be the time-t present value of the cash-flow stream generated by a worker i currently matched to firm j under the employment contract $\mathscr{C}_{i,j,\tau}$. Dropping the explicit dependence on $\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t\right)$, the value of the employment contract $\mathscr{C}_{i,j,\tau}$ to a firm in period $t \geq \tau$ can be recursively constructed using

$$V_t\left(\mathscr{C}_{i,j,\tau}\right) = y_{i,t}\left[1 - A\left(\theta_{i,j,t}\right)\right] - C_{i,j,t} + \kappa \mathbb{E}_t\left[\frac{\Lambda_{t+1}}{\Lambda_t}\theta_{i,j,t}\delta_{i,j,t+1}V_{t+1}\left(\mathscr{C}_{i,j,\tau}\right)\right].$$

In the above equation, the flow profit for the firm equals the output of the worker net of compensation and retention cost. In the next period, the match continues with probability $\kappa\theta_{i,j,t}$, and $\delta_{i,j,t+1}$ is the indicator function for the decision to voluntarily terminate. The future cash flows are discounted using state prices $\{\Lambda_t\}$.

Worker utility Let $U^*\left(h, g^t\right)$ be the highest utility a worker with human capital h can achieve after receiving an opportunity to match with a firm at aggregate history g^t. The utility for an unemployed worker i with human capital $h_{i,t}$ at time t, denoted by $\overline{U}\left(h_{i,t}, g^t\right)$, is recursively constructed using

$$\overline{U}\left(h_{i,t}, g^t\right) = \left[(1-\beta)\left(by_{i,t}\right)^{1-\frac{1}{\psi}} + \beta\overline{\mathbb{M}}\left(h_{i,t}, g^t\right)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{4}$$

where $by_{i,t}$ is the flow consumption from unemployment benefits provided by the government.[3] The term $\overline{\mathbb{M}}\left(h_{i,t}, g^t\right)$ is the certainty equivalent of the next-period utility: with probability $1-\chi$, unemployed workers stay unemployed with continuation utility $\overline{U}\left(h_{i,t+1}, g^{t+1}\right)$, and with probability χ, they are matched with a new firm and receive utility $U^*\left(h_{i,t+1}, g^{t+1}\right)$. Combining both of these possibilities,

$$\overline{\mathbb{M}}\left(h_{i,t}, g^t\right) = \left(\kappa\mathbb{E}_t\left[(1-\chi)\overline{U}\left(h_{i,t+1}, g^{t+1}\right)^{1-\gamma} + \chi U^*\left(h_{i,t+1}, g^{t+1}\right)^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}.$$

[3]See appendix A for a discussion of why an unemployed worker cannot obtain any insurance, from either the previous employer or any other firm and, as a result, consumes only the value of unemployment benefits.

For $t \geq \tau$, let $U_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}_{i,j,\tau}\right)$ be the utility of a matched worker i at time t under the employment contract $\mathscr{C}_{i,j,\tau}$. It satisfies the recursion

$$U_t\left(\mathscr{C}_{i,j,\tau}\right) = \left[(1-\beta)\left(C_{i,j,t}\right)^{1-\frac{1}{\psi}} + \beta \mathbb{M}_t\left(\mathscr{C}_{i,j,\tau}\right)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{5}$$

where

$$\mathbb{M}_t\left(\mathscr{C}_{i,j,\tau}\right) = \left(\kappa \mathbb{E}_t\left[\theta_{i,j,t}\delta_{i,j,t+1}U_{t+1}\left(\mathscr{C}_{i,j,\tau}\right)^{1-\gamma} + \left(1-\theta_{i,j,t}\delta_{i,j,t+1}\right)\overline{U}\left(h_{i,t+1}, g^{t+1}\right)^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}.$$

The computation of the certainty equivalent $\mathbb{M}_t\left(\mathscr{C}_{i,j,\tau}\right)$ accounts for the fact that the match between work i and firm j can be terminated exogenously with probability $1-\theta_{i,j,t}$ or endogenously by setting $\delta_{i,j,t+1} = 0$.

Agency frictions We impose two types of agency frictions. First, neither firms nor workers can commit. At the beginning of each period t, before production takes place, firms and workers have an opportunity to terminate their match by setting $\delta_{i,j,t} = 0$. Once a match is dissolved, the worker is unemployed, and the firm has the option of keeping open the vacancy or hiring a new worker. Second, firms' choices of effort $\theta_{i,j,t}$, which determine the probability that the match will continue to the next period, are observable neither to workers nor to any other firms. We show later that our specification of the moral hazard problem provides a tractable way to generate equilibrium separations and non-trivial labor market dynamics.[4]

The presence of agency frictions imposes incentive compatibility constraints on the feasibility of a contract $\mathscr{C}_{i,j,\tau}$. Perfect competition on the labor market and no cost for keeping or posting vacancies imply that the value of a firm's option of terminating a match is zero. Thus, incentive compatibility with respect to the firm-side limited commitment requires that the present value of any employment contract must be non-negative at all times. That is, for all match specific histories, either $V_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}_{i,j,\tau}\right) \geq 0$, or $V_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}_{i,j,\tau}\right) < 0$ and the firm voluntarily terminates the match by setting $\delta_{i,j,t}\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t\right) = 0$. Thus, for all $\left(z_{i,j}^{\tau \to t}, g^t\right)$,

$$\delta_{i,j,t} V_t\left(\mathscr{C}_{i,j,\tau}\right) \geq 0. \tag{6}$$

Similarly, a worker always has the option of terminating a match and becoming unemployed to obtain utility $\overline{U}_t\left(h_{i,t}, g^t\right)$. Therefore, the worker-side limited commitment implies that at any match-specific history $z_{i,j}^{\tau \to t}$, either the worker continues the employment relationship and obtains a utility that is higher than his outside option, or he unilaterally

[4]For a similar specification of the moral hazard problem between firms and workers, see Lamadon (2016).

terminates the match. Therefore, for all $\left(z_{i,j}^{\tau \to t}, g^t\right)$,

$$\delta_{i,j,t}\left[U_t\left(\mathscr{C}_{i,j,\tau}\right) - \overline{U}\left(h_{i,t}, g^t\right)\right] \geq 0. \tag{7}$$

Finally, the fact that retention effort is not observable implies that the choice of θ must be incentive compatible from the firm's perspective. This requires that for any match-specific history $\left(z_{i,j}^{\tau \to t}, g^t\right)$,

$$V_t\left(\mathscr{C}_{i,j,\tau}\right) \geq y_{i,t}\left[1 - A\left(\tilde{\theta}\right)\right] - C_{i,j,t} + \kappa\tilde{\theta}\mathbb{E}_t\left[\frac{\Lambda_{t+1}}{\Lambda_t}\delta_{t+1}V_{t+1}\left(\mathscr{C}_{i,j,\tau}\right)\right] \tag{8}$$

for all $\tilde{\theta}$.[5]

Feasibility and efficiency We now define feasible and privately efficient contracts. These definitions take as given a stochastic process for state prices $\left\{\Lambda_t\left(g^t\right)\right\}_t$ and values of newly employed workers $\left\{U^*\left(\cdot, g^t\right)\right\}_t$.

Definition 1. *A contract $\mathscr{C}_{i,j,\tau}$ offered by a firm to a newly employed worker with human capital $h_{i,\tau}$ in period τ is feasible given $\left\{\Lambda_t\left(g^t\right), U^*\left(\cdot, g^t\right)\right\}_t$, if it satisfies limited commitment constraints (6) and (7) and incentive compatibility constraints (8).*

Definition 2. *A contract $\mathscr{C}_{i,j,\tau}$ offered by a firm to a newly employed worker with human capital $h_{i,\tau}$ in period τ is privately efficient given $\left\{\Lambda_t\left(g^t\right), U^*\left(\cdot, g^t\right)\right\}_t$, if it is feasible, and there does not exist an alternative feasible contract $\hat{\mathscr{C}}$ such that $V_\tau\left(h, \emptyset, g^\tau | \hat{\mathscr{C}}\right) > V_\tau\left(h, \emptyset, g^\tau | \mathscr{C}_{i,j,\tau}\right)$ and $U_\tau\left(h, \emptyset, g^\tau | \hat{\mathscr{C}}\right) \geq U_\tau\left(h, \emptyset, g^\tau | \mathscr{C}_{i,j,\tau}\right)$.*

A competitive equilibrium with long-term contracts needs to specify values of newly employed workers $\left\{U^*\left(\cdot, g^t\right)\right\}_t$, and equilibrium state prices $\left\{\Lambda_t\left(g^t\right)\right\}_t$. The value of a newly matched worker $U^*\left(\cdot, g^t\right)$ is determined by workers' optimal choice of contract on the competitive labor market. Given the process $\left\{\Lambda_t\left(g^t\right)\right\}_t$, at any aggregate history g^τ, the function $U^*\left(\cdot, g^\tau\right)$ solves

$$U^*\left(h, g^\tau\right) = \max_{\mathscr{C}}\left\{U_\tau\left(h, \emptyset, g^\tau | \mathscr{C}\right) : \mathscr{C} \text{ is privately efficient given } \left\{\Lambda_t\left(g^t\right), U^*\left(\cdot, g^t\right)\right\}_t\right\}. \tag{9}$$

Because firms can always choose to keep open a vacancy, any contract offered must satisfy the condition $V_\tau\left(\mathscr{C}_{i,j,\tau}\right) \geq 0$. Among all firms that offer employment contracts, a worker chooses to match with the firm that offers the most favorable terms. An implication of equation (9) is that at the beginning of each match, competitive labor markets drive firm value to zero.

[5] We rely on the standard result in dynamic mechanism design: there is no profitable deviation in the dynamic environment if and only if one-step deviations are not profitable.

Finally, firm owners' optimal consumption and savings decisions impose a restriction on the a relationship between state prices $\left\{\Lambda_t\left(g^t\right)\right\}_t$ and firm owners' equilibrium consumption. Because the firm owners hold the equity of all firms, they are well diversified and their consumption and portfolio choices depend only on aggregate quantities. Given a stochastic process for firm-owner consumption $\left\{X_t\left(g^t\right)\right\}_t$, the utility of firm owners is given recursively by

$$W_t\left(g^t\right) = \max\left\{(1-\beta)\, X\left(g^t\right)^{1-\frac{1}{\psi}} + \beta \mathbb{N}_t\left(g^t\right)^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1-\frac{1}{\psi}}},$$

where the certainty equivalent $\mathbb{N}_t\left(g^t\right) = \left(\mathbb{E}_t W_{t+1}\left(g^{t+1}\right)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}$. As is standard in asset pricing models with recursive utility, the optimality condition for firm owner's consumption and investment choice implies that the stochastic discount factor must satisfy

$$\frac{\Lambda_{t+1}\left(g^{t+1}\right)}{\Lambda_t\left(g^t\right)} = \beta\left[\frac{X_{t+1}\left(g^{t+1}\right)}{X_t\left(g^t\right)}\right]^{-\frac{1}{\psi}}\left[\frac{W_{t+1}\left(g^{t+1}\right)}{\mathbb{N}_t\left(g^t\right)}\right]^{\frac{1}{\psi}-\gamma}. \tag{10}$$

It is important to note that our formulation does not assume any form of exogenously incomplete market. From equation (10), firm owners have access to complete markets, and their intertemporal rate of substitution is a valid stochastic discount factor. Because long-term contracts can replicate aggregate and idiosyncratic state-contingent payoffs between firm owners and workers, the same condition holds true for workers unless an incentive compatibility constraint is binding.[6] In the absence of those agency frictions, all workers in our setup have access to a complete set of state-contingent payoffs via insurance contracts, and our setup becomes isomorphic to a standard representative agent complete markets model.

In addition, a more general contract space that allows payments from all firms to all workers does not change any conclusions. To avoid cumbersome notation, we have restricted attention to simple employment contracts, where payments are described only between firms and their current employees. In Appendix A, we show that due to the two-sided limited commitment and perfect competition among firms, the simple employment contract described in (3) is in fact optimal in the larger contracting space in which firms are allowed to offer insurance contracts to unrelated workers. Intuitively, defaulting on a contract between a firm and its employer results in separation and human capital loss, which serves as commitment device to sustain limited risk sharing between employer firms and their employees. However, defaulting on a contract between an unrelated firm and worker does not exclude either of the parties from entering into insurance contracts with others. Lemmas 1 and 2 in Appendix A show that under fairly general conditions, this lack of exclusion rules out any insurance

[6] This is formally shown later in proposition 2.

between firms and unrelated workers.[7]

Equilibrium state prices, workers' outside valuations, and optimal contracts for each worker-firm pair depend on past histories of aggregate as well as firm- and worker-level idiosyncratic shocks. In general, this means that one needs to keep track of the distribution of worker characteristics such as human capital and terms of contracts within firms and across firms in the economy. This can quickly become unmanageable. However, thanks to the functional form choices on preferences and technology, we can define a competitive equilibrium with long-term contracts that is recursive in an appropriately constructed parsimonious set of state variables.

2.2 Recursive Competitive Equilibrium

To define a recursive competitive equilibrium, we follow the dynamic contracting literature (for instance, Thomas and Worrall 1988 or Atkeson and Lucas 1992) and express the contracting problem using promised utility as a state variable. We show that in our model, the homotheticity properties of preferences and technology and the random walk assumption of exogenous shocks imply that an individual worker's history can be summarized by a one-dimensional state variable u, which equals the current-period continuation utility divided by the current-period worker output. The aggregate history can be summarized by a vector of aggregate state variable $S \equiv (g, \phi, B)$. Here, g is the growth rate of aggregate productivity, ϕ is a one-dimensional measure that summarizes agent types, and B is the fraction of total output consumed by unemployed workers. That we ultimately need to keep track of only a one-dimensional distribution as a state variable is a key computational step for our quantitative analysis.

Normalized variables Before stating the normalized version of the optimal contracting problem, we define firms' value function $V(y, U, S)$ to be the maximum value of $V_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}\right)$ that can be achieved by any contract $\mathscr{C}$ such that it is feasible and provides the worker a utility of at least U; that is, $U_t\left(h_{i,\tau}, z_{i,j}^{\tau \to t}, g^t | \mathscr{C}\right) \geq U$. Homogeneity of preferences and technology implies that the value function $V(y, U, S)$ satisfies

$$V(y, U, S) = v\left(\frac{U}{y}, S\right) y \tag{11}$$

for some normalized value function v. This motivates the concept of normalized promised utility $u \equiv \frac{U}{y}$. In addition, the highest utility a worker can achieve from a new match $U^*(h, g^\tau) = u^*(S)\, y$ and the worker's utility in unemployment $\overline{U}(h, g^t) = \overline{u}(S)\, y$, where

[7] Chien and Lustig (2010) and Rampini and Viswanathan (2010, 2013) also assume a form of non-exclusion in limited commitment models.

$u^*(S)$ and $\overline{u}(S)$ are functions of aggregate states. It is possible to prove that the value function $v(u,S)$ must be strictly decreasing in u.[8] Therefore, equation (9) implies that $u^*(S)$ has to satisfy

$$u^*(S) = \max\{u : v(u,S) \geq 0\}. \tag{12}$$

In addition, (4) requires the following relationship between $\overline{u}(S)$ and $u^*(S)$:

$$\overline{u}(S) = \left[(1-\beta)\, b^{1-\frac{1}{\psi}} + \beta\lambda \overline{m}(S)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{13}$$

with $\overline{m}(S) \equiv \left(\kappa \sum_{g'} \pi(g'|g)\left[e^{(1-\gamma)g'}\left\{(1-\chi)\overline{u}(S')^{1-\gamma} + \chi u^*(S')^{1-\gamma}\right\}\right]\right)^{\frac{1}{1-\gamma}}$.

Recursive optimal contracting Let $\Gamma_\phi(g',S)$ and $\Gamma_B(g',S)$ be the laws of motion for the endogenous aggregate states ϕ and B so that $S' \equiv (g',\phi',B') = (g',\Gamma_\phi(g',S),\Gamma_B(g',S))$. Let $\{\Lambda(S',S)\}_{g'}$ be the set of one-period-ahead Arrow security prices, and let $\zeta' = (g',\eta',\varepsilon')$ be the vector of next period aggregate and match-specific shocks. The Markov transition matrix $\{\pi(g'|g)\}_{g,g'}$ together with the conditional density $f(\eta,\varepsilon|g)$ define a probability distribution for ζ' conditional on g, which we denote as $\Omega(d\zeta'|g)$.

The normalized firm value $v(u,S)$ satisfies the following Bellman equation:

$$v(u,S) = \max_{c,\theta,\{u'(\zeta'),\delta'(\zeta')\}_{\zeta'}} \begin{array}{c} 1 - c - A(\theta) + \\ \kappa\theta \int \Lambda(S',S)\, e^{g'+\eta'+\varepsilon'}\delta'(\zeta')\, v(u'(\zeta'),S')\,\Omega(d\zeta'|g), \end{array} \tag{14}$$

subject to

$$u = \left[(1-\beta)\, c^{1-\frac{1}{\psi}} + \beta m^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{15}$$

$$\delta'(\zeta')\, v(u'(\zeta'),S') \geq 0, \;\; for \;\; all \;\; \zeta', \tag{16}$$

$$\delta'(\zeta')\left[u'(\zeta') - \lambda\overline{u}(S')\right] \geq 0, \;\; for \;\; all \;\; \zeta', \tag{17}$$

$$A'(\theta) = \kappa \int \Lambda(S',S)\, e^{g'+\eta'+\varepsilon'}\delta'(\zeta')\, v(u'(\zeta'),S')\,\Omega(d\zeta'|g), \tag{18}$$

where $m = \left\{\kappa \int e^{(1-\gamma)(g'+\eta'+\varepsilon')}\left[\theta\delta'(\zeta')\, u'(\zeta')^{1-\gamma} + (1-\theta\delta'(\zeta'))\,\lambda\overline{u}(S')^{1-\gamma}\right]\Omega(d\zeta'|g)\right\}^{\frac{1}{1-\gamma}}$.

Equation (15) is the promise-keeping constraint ensuring that the current compensation and effort choices, together with the choices for future continuation values, deliver the utility u that is promised to the worker. Inequalities (16) and (17) are the recursive counterparts of the limited commitment constraints (6) and (7). Equation (18) is the first-order necessary condition for firms' choice of retention effort. Because the cost function $A(\theta)$ is strictly convex

[8] See appendix B.1 for details on existence and monotonicity of the function v.

in θ, first-order conditions (18) are equivalent to (8) and therefore, as we prove in appendix B, necessary and sufficient for incentive compatibility. We label the above maximization problem as *P1*.

Aggregation Let $x_t\left(g^t\right) = \frac{X_t(g^t)}{Y_t(g^t)}$ be the normalized consumption of the firm owners. Given a consumption function $x\left(S\right)$, firm owners' normalized utility, which we denote as $w\left(S\right)$, can be constructed from

$$w\left(S\right) = \left[\left(1-\beta\right)x\left(S\right)^{1-\frac{1}{\psi}} + \beta n\left(S\right)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{19}$$

with the certainty equivalent $n\left(S\right) = \left\{\sum_{g'} \pi\left(g'|g\right) e^{(1-\gamma)g'} w\left(S'\right)^{1-\gamma}\right\}^{\frac{1}{1-\gamma}}$. Using the normalized notation, the stochastic discount factor (SDF) $\Lambda\left(S', S\right)$ is given by

$$\Lambda\left(S', S\right) = \beta\left[\frac{x\left(S'\right)e^{g'}}{x\left(S\right)}\right]^{-\frac{1}{\psi}}\left[\frac{w\left(S'\right)e^{g'}}{n\left(S\right)}\right]^{\frac{1}{\psi}-\gamma}. \tag{20}$$

Finally, we describe the construction of the aggregate state variable ϕ, which we will refer to as the "summary measure." Let $\Phi_j\left(du, dh\right)$ denote the joint distribution of (u, h) for workers in firm j and $\Phi_0\left(dh\right)$ the distribution of human capital of unemployed workers. In general, $\left\{\Phi_j\right\}_{j=0}^{N}$ is a state variable in the construction of a recursive equilibrium because the resource constraint,

$$Y\int bh\Phi_0\left(dh\right) + Y\sum_{j=1}^{N}\int\int\left[c\left(u,S\right) + A(\theta,S)\right]h\Phi_j\left(du,dh\right) + Yx(S) = Y\sum_{j=1}^{N}\int\int h\Phi_j\left(du,dh\right), \tag{21}$$

depends on $\left\{\Phi_j\right\}_{j=0}^{N}$. Let $c\left(u,S\right)$ be the policy function for worker compensation in the problem *P1*. The total compensation to all workers

$$Y\sum_{j=1}^{N}\int\int c\left(u,S\right)h\Phi_j\left(du,dh\right) = Y\int c\left(u,S\right)\sum_{j=1}^{N}\left[\int h\Phi_j\left(dh|\,u\right)\right]\Phi_j\left(du\right),$$

where we decompose the joint distributions into a marginal distribution and a conditional distribution: $\Phi_j\left(du,dh\right) = \Phi_j\left(dh|\,u\right)\Phi_j\left(du\right)$. We define the summary measure by $\phi\left(du\right) \equiv \sum_{j=1}^{N}\int h\Phi_j\left(dh|\,u\right)$ for all u. For a given h, the term $\sum_{j=1}^{N}\Phi_j(du,dh)$ is the joint distribution of (u, h) across all firms, and thus $\phi(du)$ is the average human capital of employed workers of type u. We define the total compensation to all unemployed workers normalized by aggregate productivity as $B = \int bh\Phi_0\left(dh\right)$. Using the fact that total output equals $Y\int\phi\left(du\right)$, the resource constraint can be written as

$$B+\int\left[c\left(u,S\right)+A(\theta(u,S))\right]\phi\left(du\right)+x\left(S\right)=\int\phi\left(du\right). \tag{22}$$

The above procedure reduces the $N+1$ two-dimensional distributions $\{\Phi_j\}_{j=0}^{N}$ into a one-dimensional measure ϕ and a scalar B. This greatly simplifies our analysis.

Recursive competitive equilibrium Equilibrium can be constructed in two steps. In the first, we obtain policy functions $c\left(u,S\right)$, $\theta\left(u,S\right)$, $\{u'\left(u,S,\zeta'\right),\delta'\left(u,S,\zeta'\right)\}_{\zeta'}$ by solving problem *P1*. In the second, we use the policy functions to construct the laws of motion for the endogenous state variables Γ_ϕ and Γ_B.

The summary measure ϕ has a continuous density on $[\lambda\overline{u}\left(S\right),\ u^*\left(S\right)]$ which describes the human capital of all currently employed workers and a mass point on $u^*\left(S\right)$ for newly employed workers. The law of motion $\phi'=\Gamma_\phi\left(g',S\right)$ specifies the summary measure in the next period for each possible realization of g' as a function of current state S. The density of the continuous part of ϕ' is

$$\phi'(d\widetilde{u})=\kappa\int\theta\left(u,S\right)\left[\int e^{\eta'+\varepsilon'}f\left(\eta',\varepsilon'|g'\right)\delta\left(\zeta'\right)I_{\{u'(u,S,\zeta')\in d\tilde{u}\}}d\eta'd\varepsilon'\right]\phi\left(du\right) \tag{23}$$

$\forall\tilde{u}\in[\lambda\overline{u}\left(S'\right),\ u^*\left(S'\right)]$, where I is the indicator function. The mass point of ϕ' at $u^*\left(S'\right)$ is given by

$$\phi'\left(\left\{u^*\left(S'\right)\right\}\right)=(1-\kappa)+\kappa\chi\lambda\frac{B}{b}. \tag{24}$$

In the above expression, $1-\kappa$ is the total amount of human capital of newborn workers, a measure $1-\kappa$ of whom arrive in each period with one unit of human capital. The second term is the amount of human capital of workers who will move to employment from the current unemployed pool. The term B is total unemployment benefit in the current period and $\frac{B}{b}$ is the total human capital of all unemployed workers. A fraction κ of them survive to the next period, their human capital decpreciates at rate $1-\lambda$, and a fraction χ exit the unemployment pool.

The law of motion $\Gamma_B\left(g',S\right)$ maps $\left(g',S\right)$ to B', which is the total unemployment benefit in the next period in state g' and is given by

$$B'=\kappa\lambda\left[B(1-\chi)+b\int\int\left(1-\theta\left(u,S\right)\delta\left(\zeta'\right)\right)f\left(\eta',\varepsilon'|g'\right)\phi\left(du\right)d\eta'd\varepsilon'\right], \tag{25}$$

where the first term $B(1-\chi)$ accounts for all unemployed workers in the current period who will stay unemployed in the next period, and the second term accounts for workers who transit from the currently employed pool to unemployment in the next period.

Definition 3. *A recursive competitive equilibrium consists of stochastic discount factor* $\{\Lambda(S', S)\}$*, workers' value from unemployment* $\overline{u}(S)$*, the value from a new match* $u^*(S)$*, firm values* $v(u, S)$ *and policy functions* $\left(c(u,S), \theta(u,S), \{u'(u,S,\zeta'), \delta'(u,S,\zeta')\}_{\zeta'}\right)$*, consumption share of firm owners* $x(S)$*, and laws of motion* Γ_ϕ *and* Γ_B *such that (i) the stochastic discount factor satisfies* (20); *(ii) the firm value function and the policy functions solve problem P1; (iii) the laws of motion for aggregate states* ϕ *and* B *satisfy (23), (24), and (25); (iv) values for new and unemployed workers satisfy (12) and (13); and (v) the resource constraint (22) holds.*

3 The Optimal Contract

In this section, we provide a characterization of the optimal contract by discussing the properties of policy functions to problem *P1*. The policy functions of firm retention effort $\theta(u, S)$ and the termination decision $\delta'(u, S, \zeta')$ determine the labor market dynamics in our model, while the choices of consumption $c(u, S)$ and the next-period continuation utility $u'(u, S, \zeta')$ are responsible for the model's implications on risk sharing. We start with the labor–market related policy functions.

First, there are no voluntary terminations under the optimal contract. In the absence of complementarity between firm and worker productivities, an adverse shock to human capital makes the worker equally unproductive in all firms and proportionally lowers his eligible unemployment benefit. Separations, which lead to human capital losses, therefore lower worker utility without benefiting firms. The optimal contract avoids such inefficient separations by setting $\delta(u, S, \zeta') = 1$ for all ζ'. In our setup, although the possibility of separation serves as a punishment device and sustains some risk sharing between firms and their employees, it is never specified as an equilibrium outcome under the optimal contract.

Next, firms' retention policy function $\theta(u, S)$ is decreasing in u. Incentive compatibility constraint (18) requires that the marginal cost $A'(\theta)$ of retaining the worker equal its marginal benefit, the present value of the cash flow that the worker brings to the firm, $\kappa \int \Lambda(S', S) e^{g'+\eta'+\varepsilon'} v(u'(u, S, \zeta'), S') \Omega(d\zeta'|g)$. Because firm value is a decreasing function of the utility promised to workers, these marginal benefits are also decreasing in u. Thus, it is harder to induce a higher retention effort when the promised utility to worker is high. The following lemma summarizes the above discussion of $\delta(u, S, \zeta')$ and $\theta(u, S)$.

Proposition 1. *In any equilibrium in which the stochastic discount factor and the law of motion for aggregate state variables satisfy condition (6) in appendix B, for all* (u, S, ζ'), $\delta'(u, S, \zeta') = 1$*, and the policy function for retention effort,* $\theta(u, S)$ *is decreasing in* u *for all* S*.*

Proof. See appendix B. □

More generally, the above proposition implies that separation rates are higher when the value of a worker to the firm is low. This may be due to either a lower future surplus from the worker (that is, lower levels of $v\left(u'\left(u,S,\zeta'\right),S'\right)$) or a higher discount rate (that is, lower values of Λ). Therefore, the specification of the moral hazard problem—in particular, the incentive compatibility constraint (18)—generates countercyclical unemployment. Also, the implication that $\delta'\left(u,S,\zeta'\right)=1$ is useful for the tractability of the model. It allows us to replace the firm- and work-side limited commitment constraints in equations (16) and (17) by

$$v\left(u'\left(\zeta'\right),S'\right)\geq 0, \tag{26}$$

$$u'(\zeta')\geq\lambda\overline{u}\left(S'\right). \tag{27}$$

We now turn to the implications of the optimal contract for risk sharing. With full commitment, firms can perfectly insure workers against idiosyncratic shocks. Therefore, workers' continuation utilities $ye^{\varepsilon'+\eta'}u'\left(u,S,\zeta'\right)$ do not respond to these shocks and are equalized across all possible realizations of (η',ε'). When 0 is a possible realization of η' and ε', this optimal risk-sharing condition can be written as

$$u'\left(u,S,\zeta'\right)=e^{-(\varepsilon'+\eta')}u'\left(u,S,g',0,0\right),\forall\zeta'. \tag{28}$$

Thus, under perfect risk sharing, the elasticity of normalized utility with respect to idiosyncratic shocks is -1.

Under limited commitment, equation (28) cannot hold for all values of (η',ε'). Because a worker can always separate voluntarily, the promised utility under the optimal contract cannot be lower than what he receives upon a voluntary termination, $\lambda\overline{u}\left(S'\right)$. Clearly, for large and positive realizations of $\varepsilon'+\eta'$, the full risk sharing policy in (28) would imply that $e^{-(\varepsilon'+\eta')}u'\left(u,S,g',0,0\right)<\lambda\overline{u}\left(S'\right)$ and violate the worker-side limited commitment (27). As a result, given the current state (u,S), there is a threshold level $\overline{\varepsilon}(u,S,g')$ for every g', such that for all $\varepsilon'+\eta'\geq\overline{\varepsilon}(u,S,g')$, the worker-side limited commitment constraint binds and

$$u'\left(u,S,\zeta'\right)=\lambda\overline{u}\left(S'\right). \tag{29}$$

Conversely, the firm-side limited commitment imposes an upper bound on $u'\left(u,S,\zeta'\right)$. Because $u^*\left(S'\right)$ is the highest utility a worker can achieve and $v\left(u^*\left(S'\right),S'\right)=0$, any promised utility higher than $u^*\left(S'\right)$ results in a negative firm value. Large and negative realizations of $\varepsilon'+\eta'$ therefore imply that the full risk sharing policy (28) would lead to $e^{-(\varepsilon'+\eta')}u'\left(u,S,g',0,0\right)>u^*\left(S'\right)$ and thus violate the firm-side limited commitment

(26). Under the optimal contract, there is a threshold function $\underline{\varepsilon}(u, S, g')$, such that for all $\varepsilon' + \eta' < \underline{\varepsilon}(u, S, g')$, the firm-side limited commitment constraint has to bind, and

$$u'\left(u, S, \zeta'\right) = u^*\left(S'\right). \tag{30}$$

In the interior, $\underline{\varepsilon}(u, S, g') < \varepsilon' + \eta' < \overline{\varepsilon}(u, S, g')$, none of the above constraints bind, and the intertemporal marginal rate of substitution of all agents has to equalize. In the following proposition, we summarize the properties of consumption and continuation utility policies.

Proposition 2. *In any equilibrium in which the stochastic discount factor and the law of motion for aggregate state variables satisfy condition (6) in appendix B, there exist threshold levels $\underline{\varepsilon}(u, S, g')$ and $\overline{\varepsilon}(u, S, g')$ with $\underline{\varepsilon}(u, S, g') < \overline{\varepsilon}(u, S, g')$, such that for all $\varepsilon' + \eta' > \overline{\varepsilon}(u, S, g')$, $u'(u, S, \zeta')$ is given by (29), and for all $\varepsilon' + \eta' < \underline{\varepsilon}(u, S, g')$, $u'\left(u, S, \zeta'\right)$ satisfies (30). For all $\varepsilon' + \eta' \in [\underline{\varepsilon}(u, S, g'), \overline{\varepsilon}(u, S, g')]$, $u'\left(u, S, \zeta'\right)$ is strictly decreasing in $\varepsilon' + \eta'$ and satisfies*

$$\left[\frac{x\left(S'\right)}{x\left(S\right)}\right]^{-\frac{1}{\psi}}\left[\frac{w\left(S'\right)}{n\left(S\right)}\right]^{\frac{1}{\psi}-\gamma}\left(1+\frac{\iota\left(u,S\right)}{\theta\left(u,S\right)}\right) = e^{-\gamma(\eta'+\varepsilon')}\left[\frac{c\left(u'\left(u,S,\zeta'\right),S'\right)}{c\left(u,S\right)}\right]^{-\frac{1}{\psi}}\left[\frac{u'\left(u,S,\zeta'\right)}{m\left(u,S\right)}\right]^{\frac{1}{\psi}-\gamma}, \tag{31}$$

where $\iota\left(u, S\right) > 0$ is given in appendix B.

Proof. See appendix B. $\square$

The above proposition has several implications. First, large and positive realizations of ε' and η' imply that $u'\left(u, S, \zeta'\right)$ must be set to a constant and cannot respond to further increases in $\eta' + \varepsilon'$. As a result, the level of continuation utility, $ye^{\varepsilon'+\eta'}u'\left(u, S, \zeta'\right)$, must increase with positive productivity shocks. High promised values are met with higher future compensation. This feature of our setting is similar to that in Harris and Holmstrom (1982), Kehoe and Levine (1993), and Alvarez and Jermann (2000).

Second, in contrast to these papers in which workers are perfectly insured against downside risk, the limited commitment constraint on the firm side implies that sufficiently negative realizations of $\eta' + \varepsilon'$ also cannot be hedged. A sequence of negative worker- or firm-specific productivity shocks lowers worker output. Keeping an extremely unproductive worker is a negative net present value undertaking for the firm, since the cash flow produced by the worker is not enough to pay for his promised compensation. In addition to lower retention effort as mentioned above, lack of commitment from the firm side requires reductions in future worker compensation in order to provide incentives for the firm to continue the match. As we will demonstrate in subsequent sections, this feature is key for our model to generate volatile asset prices along with tail risk in labor earnings.

Third, equation (31) in proposition 2 implies that the intertemporal marginal rate of substitution has to be equal for all agents in the economy unless the limited commitment constraints are binding. This includes firm owners as well as a subset of workers. Equation (31) also highlights the difficulty in generating a high volatility of stochastic discount factor without directly assuming exogenous market segmentation. Given preference parameters, the volatility of the SDF is determined entirely by the risk exposure of the consumption of marginal investors. If a heterogeneous agent-based model such as ours generates a stochastic discount factor that is more volatile than a representative agent model, then it must also provide an explanation for why the additional risk exposure in marginal investors' consumption is not insured away through conditions like equation (31). In the next section, we demonstrate conditions under which agency frictions generates downside risk in labor earnings and amplify the volatility of the stochastic discount factor in equilibrium.

4 Agency Frictions and Asset Pricing

In this section, we highlight how agency frictions affect aggregate and cross-sectional asset returns. General equilibrium linkages between tail risk in labor earnings and the pricing kernel are key for agency frictions to amplify risk premia. We start with an "irrelevance" result in the spirit of Krueger and Lustig (2010) that provides conditions under which agency frictions are irrelevant for both the price of aggregate risk and aggregate labor market dynamics. We then analyze a special case of our model to isolate the mechanism that amplifies the volatility of the stochastic discount factor and distinguish it from alternatives in the literature. We also derive a set of testable predictions of our model mechanism, which are later confronted with the data.

4.1 An Irrelevance Result

Krueger and Lustig (2010) show that if the aggregate endowment growth is i.i.d. and the distribution of idiosyncratic shocks $f\left(\eta,\varepsilon\middle|\, g\right)$ is independent of the aggregate states, then uninsurable idiosyncratic risk does not affect the price of aggregate shocks in a wide set of incomplete markets models. To formalize a version of their result in our setting with contracting frictions, we start with the following definition.

Definition 4. *Given the economy described in section 2.2, an equivalent deterministic economy with a modified discount rate is an otherwise identical economy except that the aggregate growth rate is set to zero and the time discount factor β is modified to $\beta\left(\mathbb{E}\left[e^{(1-\gamma)g'}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}$.*

In the following proposition, we provide conditions under which a recursive competitive equilibrium in the stochastic economy can be constructed from the equilibrium of an equivalent deterministic economy with a modified discount rate.

Proposition 3. *(Krueger and Lustig) Suppose that g_t is i.i.d. over time and $f(\eta, \varepsilon| g)$ does not depend on g. If there exists an equilibrium in the equivalent deterministic economy with a modified discount rate, then there exists an equilibrium of the stochastic economy described in section 2.2 with the stochastic discount factor*

$$\Lambda(S', S) = \frac{1}{\hat{R}(\phi, B)} \frac{e^{-\gamma g'}}{\mathbb{E}\left[e^{(1-\gamma)g'}\right]}, \tag{32}$$

where $\hat{R}(\phi, B)$ is the risk-free interest rate in the equivalent deterministic economy with a modified discount rate.

Proof. See section A in the online appendix (not for publication). □

With i.i.d aggregate growth rates, the stochastic discount factor in the section 2.2 economy with full commitment and no moral hazard equals $\beta e^{-\gamma g'}$.[9] This is also the stochastic discount factor for a representative agent economy in which the growth rate of aggregate consumption is g_t. Equation (32) states that the stochastic discount factor in the economy with agency frictions differs only by a multiplicative constant. Therefore, agency frictions affect the risk-free interest rate but are irrelevant for the pricing aggregate risks. Proposition 3 imposes no other restriction on the distribution of idiosyncratic risk and, in particular, allows $f(\eta, \varepsilon| g)$ to contain a fat tail as long as it is the same across all realizations of g. We show in appendix B that the optimal contract in the equivalent deterministic economy with a modified discount rate can be used to construct the optimal contract in the stochastic economy by simply adjusting for aggregate growth and that the consumption share of firm owners in the stochastic economy equals that in the equivalent deterministic economy.

4.2 Aggregate Implications

Proposition 3 tells us that for agency frictions to have an impact on aggregate risk premia, we must deviate from its assumptions of i.i.d. growth and a time-invariant distribution of idiosyncratic shocks. In the rest of this section, we use a special case of our model to analyze such a departure. The special case highlights the interaction between agency frictions, risk in labor earnings, and the market price of aggregate risk.

We proceed by making several simplifying assumptions. Many of these assumptions are designed to isolate features and implications that are novel to our setting and to help us

[9]In absence of agency frictions, the consumption share of firm owners $x(S)$ is constant, and equations (19) and (20) simplify to $\Lambda(S', S) = \beta e^{-\gamma g'}$.

obtain closed-form solutions for equilibrium returns. We relax these assumptions later in the quantitative section, where we use numerical methods to solve the general model described in section 2.2.

Assumption 1. *Aggregate shocks* $g_t \in \{g_L, g_H\}$ *with* $g_L < g_H$. *From period one on, the transition probability from state* g *to state* g' *satisfies* $\pi(g'|\, g) = 1$ *if* $g' = g$. *Each firm has a single worker and* $\eta = 0$. *Let the distribution* $f(\varepsilon|g = g_H)$ *be degenerate, and the distribution* $f(\varepsilon|g = g_L)$ *be a negative exponential with parameter* ξ.[10]

This assumption includes the main departures from proposition 3. To capture the persistence of aggregate shocks, we assume that booms ($g_t = g_H$) and recessions ($g_t = g_L$) are permanent. To parsimoniously model countercyclical idiosyncratic shocks, we impose no idiosyncratic shocks in booms and negatively exponentially distributed shocks in recessions. The assumption that firm-level shocks $\eta = 0$ is without loss of generality, since proposition 2 shows that the optimal contract depends only on $\varepsilon + \eta$. In what follows, we interpret ε as both a firm- and a worker-level shock.

Assumption 2. *Preferences satisfy* $\gamma \geq \psi = 1$.

The crucial part here is that $\gamma \geq \psi$. The assumption of unit elasticity of intertemporal substitution is merely for tractability.

Assumption 3. *Workers can fully commit.*

As shown in proposition 2, uninsurable risk in the left tail of labor earnings comes from the firm-side limited commitment and separations. In section 6.1, we show the that worker-side limited commitment has little quantitative impact on the equity premium but matters for accounting for patterns in earning dynamics. Hence, here we abstract from the lack of commitment on the worker side.

Assumption 4. *Effort is costly only in period one, in which case* $A(\theta) = a\left[\ln\left(\frac{1}{1-\theta}\right) - \theta\right]$ *for some* $a > 0$.

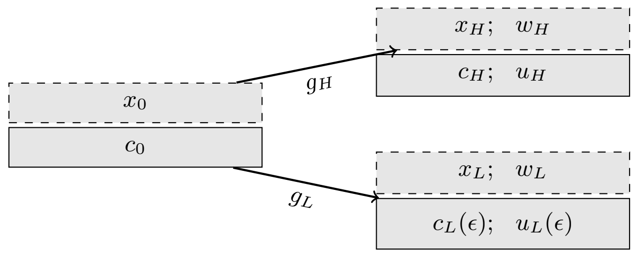
The parameter a in function $A(\theta)$ measures the severity of the moral hazard problem, with $a = 0$ corresponding to the case in which effort is costless and moral hazard is irrelevant.

Assumption 5. *For* $t = 2, 3, \ldots$, *both employed and unemployed workers produce output and consume* α *fraction of their output:* $C_t = \alpha y_t$.

From period 2 on, workers consume a fixed fraction of their output. This assumption captures that because of lack of full risk sharing, workers' consumption exposed to

[10] The form of the negative exponential distribution is described in appendix C.

Figure 1: Timing of the simple model



idiosyncratic shocks in future recessions. We assume that unemployed workers lose $1 - \lambda$ fraction of their human capital but keep producing output. They are otherwise subject to the same law of motion of human capital as employed workers.

In figure 1, we plot an event tree for the simple economy. Let the firm owners' consumption share in period 0 be x_0, and let workers' initial promised utility be u_0. We assume all workers have the same promised utility u_0; therefore, there is a unique u_0^* that is consistent with the aggregate resource constraint. In comparative static exercises, we study optimal contracting with an arbitrary u_0, even though in equilibrium the measure of agents at u_0 might be zero unless $u_0 = u_0^*$. We let $x_H \equiv x(g_H)$ and $x_L \equiv x(g_L)$, and let $w_H \equiv w(g_H)$ and $w_L \equiv w(g_L)$. For an arbitrary initial promised utility u_0, we use $\theta_H(u_0) \equiv \theta(u'(u_0, g_H), g_H)$ and $\theta_L(u_0, \varepsilon) \equiv \theta(u'(u_0, g_L, \varepsilon), g_L)$ to denote the effort choice and likewise for compensation policy $\{c_H(u_0), c_L(u_0, \varepsilon)\}$, $\{v_H(u_0), v_L(u_0, \varepsilon)\}$ for firms' value function at nodes H and L, respectively. The policy functions for compensation, firm effort, and the value functions at node H do not depend on ε since there is no idiosyncratic shock at node H. The following proposition provides conditions under which agency frictions amplify the equity premium and generate countercyclical unemployment.

Proposition 4. *(Aggregate Implications) Under assumptions 1–5, for expected utility preferences, i.e., $\gamma = 1$, firm owners' consumption share is countercyclical; that is, $x_H < x_L$. For general recursive utility, there exists a $\hat{\gamma} \in [1, 1 + \xi)$ such that if $\gamma > \hat{\gamma}$, then (i) firm owners' consumption share is procyclical; that is, $x_H > x_L$ and (ii) separation rates are countercyclical; that is, $\theta_H(u_0) > \theta_L(u_0, \varepsilon)$ for all (u_0, ε).*

Because the consumption Euler equation must hold for the unconstrained firm owners, amplification in the market price of risk relative to a representative agent model is equivalent to firm owner's consumption share being procyclical. The first part of proposition 4 implies that countercylical idiosyncratic risk by itself is not sufficient for amplifying the volatility of the equilibrium stochastic discount factor. Independent of the risk aversion γ, the optimal contract generates uninsurable tail risk (proposition 2). However, under expected utility, the pricing kernel is less volatile than the pricing kernel in an otherwise identical economy with full commitment.

Countercyclical idiosyncratic risk means that relative to booms, a larger fraction of worker-firm pairs are constrained in recessions. Because constrained firms cut compensation, in the aggregate, there is a higher fraction of resources available for firm owners during a recession. Since goods markets need to clear, these resources are allocated between the firm owners and the unconstrained workers by equating their intertemporal marginal rates of substitution. With expected utility, this amounts to equalizing the growth rates of consumption of the firm owners and the unconstrained workers. Therefore, for $\gamma = 1 = \frac{1}{\psi}$, the consumption share for firm owners increases in a recession, resulting in $x_L > x_H$. This makes aggregate asset prices less volatile.

The second implication of proposition 4 is that keeping the intertemporal elasticity of substitution fixed, a large enough risk aversion results in a procyclical consumption share for firm owners. Why does the cyclicality of x flip signs when risk aversion becomes larger relative to inverse of the intertemporal elasticity of substitution? Persistent recessions that are associated with a lack of risk sharing in the future imply lower continuation values for all workers. As risk aversion exceeds the inverse of the intertemporal elasticity of substitution, contemporaneous marginal utilities are decreasing functions of continuation utility.[11] Optimal risk sharing, which requires equating marginal rates of substitution between firm owners and unconstrained workers, is now achieved by transferring resources away from the firm owners in recessions. Proposition 4 says that for sufficiently high risk aversion, this incentive is strong enough to dominate the effect of market clearing and delivers procylical consumption shares for firm owners.

The last part of proposition 4 says that separation rates are higher in recessions relative to booms. In our model, labor income has two sources of tail risk. First, the distribution of productivity shock ε has a left tail. As shown in proposition 2, under firm-side limited commitment, this tail risk cannot be fully insured within optimal labor compensation contracts. Second, workers become unemployed with probability θ and lose a fraction $1-\lambda$ of human capital in each period until they are matched with a new firm.

The countercyclicality of unemployment risk asserted in part (ii) of proposition 4 is a direct consequence of incentive compatibility under moral hazard. Without moral hazard, to efficiently deliver promised utility to workers, firms will typically choose a lower separation rate in recessions, when human capital depreciation and consumption reduction are more costly for workers. With moral hazard, such arrangements are no longer incentive compatible,

[11] Ai and Bansal (2018) define the class of preferences under which marginal utility decreases with continuation utility as generalized risk-sensitive preferences. Generalized risk sensitivity is the key property of preferences captured by the assumption $\gamma > \frac{1}{\psi}$ that is responsible for the procyclical consumption share in our model.

and firms equalize the marginal cost of retention effort to the present value of profits that a worker can create without considering the cost of separation to workers. Since valuation ratios in recessions are lower relative to booms, firms exert less effort to retain workers, leading to countercyclical separation rates.

The effects of limited commitment and those of moral hazard reinforce each other to amplify the volatility of the stochastic discount factor. Limited commitment amplifies risk prices because optimal contracts insure workers against adverse aggregate shocks which makes firm owners' consumption more risky. Higher separations in recessions magnify the downside risk in labor earnings and hence the need for insurance. Thus, higher separation risk leads to more procyclical consumption for marginal agents, and the resulting higher discounting in turn leads to lower worker valuations, lower retention effort from firms, and more separations.

Contrasting the mechanism to alternatives proposed in the literature Proposition 4 contrasts our setup with several exogenously incomplete market models—for example, Constantinides and Duffie (1996), Constantinides and Ghosh (2014), and Schmidt (2015), as well as setups that impose exogenous market segmentation such as Basak and Cuoco (1998) or Guvenen (2009). In these papers, agents are not allowed to offer any insurance contracts to one another. In our model, agents are allowed access to a rich set of state-contingent payoffs, the only restriction being incentive compatibility constraints. In the Constantinides and Duffie (1996)–style models, all agents are marginal investors in risky assets, and hence countercyclical uninsurable risk in consumption automatically translates into a more volatile pricing kernel. In our simple example, because market incompleteness is determined by agency frictions, agents with adverse idiosyncratic shocks are constrained and therefore are not marginal. Hence, higher idiosyncratic volatility by itself does not increase the market price of risk.

Alvarez and Jermann (2001) and Chien and Lustig (2010) derive asset pricing implications in a setting with one-sided limited commitment. This corresponds to a version of our model where firms can fully commit but workers cannot. Such environments produce high equity premia when more workers are constrained in adverse aggregate states. The worker-side limited commitment constraint binds for worker-firm pairs that receive large positive idiosyncratic productivity shocks. Because constrained workers need to be compensated with higher current consumption, by market clearing, the consumption for unconstrained agents drops, raising their marginal utilities. To amplify the risk premium, such a model would necessarily require more positive skewness in labor earnings in recessions relative to booms; an implication that is inconsistent with the key feature of labor market risk that we highlight in the introduction. In addition, quantitatively, uninsurable tail risk on the downside are much more powerful than on the upside in amplifying the volatility of the stochastic discount factor. The workings of the simple example explain how a combination of firm-side limited commitment with recursive utility jointly deliver downside risk in labor earnings and higher risk premia.

Proposition 4 also distinguishes our model from Favilukis and Lin (2016b), and other papers that use sticky wages to explain the high equity premium. In these models, markets are complete and labor compensation contracts do not affect the pricing kernel. These models produce higher equity premia through an "operating leverage" channel: labor compensation is less sensitive to aggregate shocks, and this amplifies the risk exposure of capital income. Since operating leverage affects only the volatility of cash flows, these models need to assume a high level of risk aversion to match aggregate Sharpe ratios. In contrast to models with exogenous wage rigidity, in our setup, risk premia are amplified primarily through the effect of agency frictions on the volatility of the stochastic discount factor and not because of a higher volatility of dividends.[12] We return to this implication in our quantitative analysis in section 6.1.

4.3 Cross-Sectional Implications

In addition to the implications for aggregate risk prices and aggregate unemployment dynamics, our model has predictions for the cross section of returns and labor earnings. We outline these implications here, and in section 6.3 we formally test them using panel data on firm-level returns and firm-level labor shares.

In our model, heterogeneity in firms is summarized by a single state variable u. High-u firms promise a larger fraction of current and future cash flow to workers than low-u firms. Thus, u can be interpreted as "labor leverage." For a firm with labor share u_0, define the elasticity of wage payments with respect to idiosyncratic shocks as $\Upsilon(u_0) = \mathbb{E}\left[\frac{\partial \ln[e^{\varepsilon} c_L(u_0,\varepsilon)]}{\partial \varepsilon}\right]$. The term $e^{\varepsilon} c_L(u_0,\varepsilon)$ is the level of compensation to a worker with initial promised utility u_0 at node L. Next, define the valuation risk exposure or beta of a firm indexed by u_0 as $\mathcal{B}(u_0) = \left(\frac{v_H(u_0)}{\mathbb{E}[e^{\varepsilon} v_L(u_0,\varepsilon)]}\right)$. Below, we provide two comparative static results with respect to u_0.

Proposition 5. *(Cross-Sectional Implications) Under assumptions 1–5, (i)* $\frac{\partial}{\partial u_0}\Upsilon(u_0) > 0$ *and (ii) there exists a* $\hat{\gamma} \in [1, 1+\xi)$ *such that* $\forall \gamma > \hat{\gamma}$, $\exists \hat{u}$, *where* $\hat{u}$ *is defined by* $\underline{\varepsilon}(\hat{u}, g_L) = \ln \frac{1+\xi}{\xi}$, *such that* $\forall u_0 < \hat{u}$, $\frac{\partial}{\partial u_0}\mathcal{B}(u_0) > 0$.

Part (i) of the proposition says that the average elasticity of compensation with respect to idiosyncratic shock ε is increasing in promised utility u_0. Firms that promised a

[12] In our model, the claim on aggregate dividends also has a higher price-to-dividend ratio in booms relative to recessions. In section B of the online appendix (not for publication), we show that under assumptions 1–5, $\exists\ \hat{\gamma} \in [1, 1+\xi)$ such that $\gamma > \hat{\gamma}$ implies $\frac{v_H(u_0^*)}{E[e^{\varepsilon} v_L(u_0^*,\varepsilon)]} > 1$.

higher fraction of cash flows to workers are more likely to be constrained. Whenever the limited commitment constraint binds, perfect risk sharing is no longer possible, and worker compensation responds to idiosyncratic productivity shocks. Thus, labor shares predict firm-level wage pass-throughs. In section 6.3, we show that consistent with the above implication of our model, payments to workers in firms with higher labor leverage are more sensitive to firm-level idiosyncratic shocks.

Part (ii) of the proposition summarizes our model's implications for the cross section of equity returns. Compensation contracts insure workers against aggregate shocks, which makes the residual dividends more risky. In our model, firms with high u_0 have low market-to-book ratios and high labor leverage. In the cross section, the operating leverage effect is stronger for high u_0 firms. These firms promise a large fraction of their cash flow to workers, bear more aggregate risk, and compensate investors by delivering higher expected returns. In section 6.2, we use panel data on firm-level measures of labor obligations and equity prices to show that low market-to-book ratio and high labor leverage firms indeed have higher expected returns.

From a worker's perspective, proposition 5 implies that the exposure of their consumption is increasing in promised utility, which is a measure of their wealth. This implication further contrasts our setup with those with only worker-side limited commitment. In these settings, (rich) agents who have experienced a history of positive shocks are more likely to be constrained, and (poor) agents who have experienced a history of negative shocks are typically not constrained. For these models to generate an amplified equity premium, the marginal rate of substitution of the unconstrained agents necessarily needs to be more volatile than that of an average agent. Therefore, they imply that the unconstrained poor agents' risk exposure to the stock market must be higher than that of the constrained wealthy agents, an implication that is inconsistent with the empirical evidence on stockholding patterns by wealth and income.[13]

5 Quantitative Analysis

5.1 Numerical Algorithm

Policy functions and state prices depend on the infinite-dimensional state variable ϕ. The distribution ϕ shows up directly in the market clearing condition and indirectly as an argument in the stochastic discount factor when we describe the optimal contracting problem *P1*. We use a numerical procedure similar to that in Krusell and Smith (1998) and replace the distribution ϕ with suitable summary statistics. We assume that agents compute future

[13]See for instance Malloy et al. (2009).

state prices by projecting the stochastic discount factor on the space spanned by $\{g_t, x_t\}$ and use $x_{t+1} = \Gamma_x(x_t, g_t, g_{t+1})$ as a forecasting rule for x_t. Our choice of the forecasting rule is numerically efficient because given a law of motion for x, the stochastic discount factor is completely pinned down.[14]

Using the forecasting function Γ_x, we compute the stochastic discount factor $\Lambda(g', x, g)$. With $\Gamma_x(x, g, x')$ and $\Lambda(g', x, g)$, we solve the Bellman equation for the optimal contracting problem using an endogenous grid method and value function iteration. In appendix D, we describe a procedure that uses a grid on $\underline{\varepsilon}(u, S, g')$, which is the threshold for the idiosyncratic shock such that the firm-side limited commitment constraint binds, to tractably solve the contracting problem *P1*. After approximating the policy functions, we simulate a panel of agents and use the simulated data to update the law of motion Γ_x. We repeat this procedure until the function Γ_x converges. Appendix D describes the detailed steps and related diagnostics.

5.2 Calibration

Model parameters are divided into two sets: (i) parameters governing the stochastic process for aggregate shocks and (ii) parameters governing labor market flows and the distribution of idiosyncratic shocks to workers' human capital.

Aggregate shocks A period is a quarter. We time aggregate outcomes and report annual moments. We assume that the aggregate productivity process $\{g_t\}_t$ is a sum of a two-state Markov chain and a homoskedastic i.i.d. Gaussian component:[15]

$$\ln Y_{t+1} - \ln Y_t = g_{t+1} + \sigma_{\mathcal{E}} \mathcal{E}_t.$$

The state space for the Markov chain is $\{g_H, g_L\}$. We refer to states with $g = g_H$ as "booms" and states with $g = g_L$ as "recessions." The aggregate shock process $\{g_t, \mathcal{E}_t\}_t$ is calibrated as in Ai and Kiku (2013). They jointly estimate the values for $\{g_H,\, g_L\}$, the Markov transition matrix, and the volatility parameter $\sigma_{\mathcal{E}}$ from post-war aggregate consumption data. Our calibration implies an average duration of 12 years for booms and four years for recessions. The parameters for aggregate shocks are listed in the top part of table 1.

[14]The market clearing condition equation (22) implies that x is linear in a $c(u, S)$ weighted average of the distribution ϕ. It summarizes information in ϕ by assigning relatively more weight to values of u that have a larger effect on aggregate resources. This choice contrasts our algorithm to that in Krusell and Smith (1998), who use the first moment of the distribution of wealth as a summary statistic.

[15]To better fit the autocorrelation of aggregate consumption growth, we use a more flexible process than the one listed section 2.1. Equilibrium prices and the optimal contract satisfy a homogeneity property, and the presence of i.i.d $\mathcal{E}$ shocks does not increase the state space for the value and policy functions.

Labor market flows and evolution of human capital We calibrate the parameters that govern labor market flows and the evolution of human capital using transition rates between employment status, estimates of earning losses after separation, cross-sectional moments of labor earnings distributions, and other aggregate moments such as the mean and volatility of total labor compensation relative to aggregate consumption. Below, we specify our functional form choices and discuss the identification of key parameters by pairing them with the most relevant moments.

We set $\kappa = 1\%$ to obtain an average working life of 25 years. To better match the observed aggregate unemployment dynamics, we use a richer specification for the cost of retention function $A(\theta, g) = e^{a_0[\theta - a_{1,g}]}$. The scale of the cost function is normalized such that overall the costs are negligible relative to the total output. We interpret a separation in the model as a transition to the state of long-term unemployment (12 months and beyond). The parameters for $A(\cdot)$, and $\{\chi, \lambda, b\}$ are pinned down by the transition rates from employment to long-term unemployment in booms and recessions, the duration of long-term unemployment, the average earnings losses upon separation, and the estimate of the flow value of unemployment. To compute the flows in and out of long-term unemployment, we use data from the Current Population Survey summarized in table 1 of Shibata (2015). For earnings losses on separation, we use information from Davis and von Wachter (2011), who estimate the present value of earning losses due to job separations. We target the consumption equivalent of the flow value of unemployment to be 65% of pre-separation wage earnings.[16] The parameters and moments related to labor flows are listed in the middle panel of table 1.

Workers' human capital is affected by worker- and firm-level idiosyncratic shocks $\varepsilon + \eta$. We assume $\varepsilon = \alpha\varepsilon_W$ and $\eta = (1-\alpha)\varepsilon_F$, where ε_W and ε_F are i.i.d. according to a continuous density $f(\cdot|g)$. To capture the feature that the (negative) skewness of labor earnings is cyclical, we model the distribution $f(\cdot|g)$ to be a Gaussian distribution in booms and a mixture distribution of a Gaussian and a fat-tailed distribution with a negative exponential form in recessions. We assume that both the Gaussian distributions as well as the negative exponential distribution satisfy a normalization such that the exponential of the draw has a unit mean. These restrictions imply $f(\cdot|g)$ is paramterized by the following: the standard deviation of the Gaussian distribution for booms σ_H; the standard deviation of the Gaussian distribution for recessions σ_L; the intensity parameter for the negative exponential distribution ξ; and the mixture weight $\rho \in (0, 1)$, which is the probability of drawing from the negative exponential distribution in recessions.

[16]The empirical labor literature has a wide range of values for the flow value of unemployment. Shimer (2008) uses the unemployment insurance replacement rate of 40%, Rudanko (2011), and Mulligan (2012) add the value of home production and leisure and target a higher number of 85%, and Hagedorn and Manovskii (2008) use an even higher estimate of about 95%.

We set the parameter α to match the within- and across-firm variations in labor earnings as reported in Song et al. (2015) and calibrate the parameters $\{\sigma_H, \sigma_L, \rho, \xi\}$ to match the cyclical properties of the moments of labor earnings calculated using the Panel Study of Income Dynamics (PSID).[17] We restrict the sample to households where the "head of household" is a male whose working age is between 15 and 64 and who reports at least 500 hours of work in two consecutive years. Our measure of earnings is the regression residual of post-tax labor earnings on observable characteristics: age of the head, the age square, family size, and education level of the head. To obtain our target moments, we compute the cross-sectional standard deviation and Kelly skewness for log earnings growth, which are then averaged separately for "boom years" and "recession years."[18] In the bottom part of table 1, we report the parameter values and moments related to the earnings distribution.

Our model closely matches the standard deviations of the earnings growth in booms and recessions. We obtain a Kelly skewness of -3% in booms and -10% in recessions, as compared with -3.2% and -9%, respectively, in the PSID.[19]

All parameters affect the aggregate labor share. In our model, the employed workers' consumption as a fraction of aggregate consumption is countercyclical. It has a mean of 70%, a standard deviation of 3%, and an autocorrelation of 0.58. These moments are consistent with the data of aggregate labor compensation. We use national income and product accounts (NIPA) to compute the ratio of aggregate labor compensation to aggregate consumption and then detrend the series. For the sample 1947–2015, the mean labor share in consumption is 75%, the standard deviation is 2.94%, and the autocorrelation is 0.88.

6 Results

We discuss the implications for asset pricing and labor market dynamics.

6.1 Aggregate Asset Prices

In table 2, we summarize aggregate asset pricing moments. The baseline calibration is under the column labeled "Model-Baseline," and the column labeled "Model-No Frictions" is the version without limited commitment and moral hazard. We report the properties of returns

[17]The PSID is a longitudinal household survey of U.S. households with a nationally representative sample of over 18,000 individuals. Information on these individuals and their descendants has been collected continually, including data covering employment, income, wealth, expenditures, health, education, and numerous other topics. The PSID data were collected annually during the period 1968–97 and biennially after 1997.

[18]We treat 1980–82, 1991–92, 2000–01, and 2007–09 as recession years and the remaining ones as boom years.

[19]In a previous version, we also reported results for an alternative calibration that targeted moments from Guvenen et al. (2014) and produced similar asset pricing results. Compared to the Guvenen et al. (2014) data, the PSID allows us to control for transfers from the government and lifecycle earning patterns that we abstract from in our setup.

Table 1: PARAMETERS

Parameters	Values	Targeted moments	Values
Aggregate Risk			
g_H, g_L	0.35%, -0.15%	Mean, std of consumption growth	1.08%, 2.14%
$\pi(g_H\|g_H)$	0.99	Duration of booms	12 yrs
$\pi(g_L\|g_L)$	0.95	Duration of recessions	4 yrs
$\sigma_{\mathcal{E}}$	1.2%	Autocorr of consumption growth	0.44
Labor Market			
$a_{1,H}, a_{1,L}$	.995,.9925	Annualized separations rates	2%, 3%
χ	8%	Long-term unemployment duration	3 years
λ	96%	PV of earning losses on separation	30%
b	1	Flow value of unemployment	40-95%
κ	0.99	Duration of working life	25 years
Idiosyncratic Risk			
α	82%	Across firm wage variation	40%
σ_L, σ_H	7.0%, 8.0%	Std. of labor earnings change in booms and recessions	32%, 31%
τ, ρ	4.155, 2%	Kelly skewness of labor earnings change in booms and recession	-3.2%, -8.9%
Other parameters			
β,ψ,γ	0.989, 2, 5	Discount factor, IES, risk aversion	

Notes: All reported moments are annualized. The NIPA sample for aggregate consumption is 1930-2007. We follow the estimation procedure in Ai and Kiku (2013). The CPS transition rates are computed using the monthly average of workers' transitions over 12-month intervals between January 1976 and July 2014. Davis and von Wachter (2011) use longitudinal Social Security records from 1974 to 2008. The earnings losses are computed using job displacements defined as in long-tenure men, 50 years or younger, in mass-layoff events at firms with at least 50 employees. The earnings losses are accumulated for 20 years at a discount rate of 5% and are expressed as a percentage of displaced workers' average annual predisplacement earnings. The flow value of unemployment is relative to wages and in the range of estimates in Shimer (2008), Rudanko (2011), and Hagedorn and Manovskii (2008). The within- and between- firm wage variation is taken from table 6 in Song et al. (2015). We use the PSID for periods 1968-2014. The sample selection is explained in the text.

on both a claim to aggregate consumption $Y_t \int \phi_t(du)$ and a claim to aggregate corporate dividends $x_t Y_t$. Our model generates a high equity premium and a low risk-free interest rate with a risk aversion $\gamma = 5$ and an IES $\psi = 2$. Without assuming any financial leverage, the equity premium on the claim to corporate dividends is about 3.67% per year in the baseline model. The average debt-to-equity ratio for publicly traded U.S. firms is about 50%.[20] Accounting for financial leverage, our model implies a market equity premium of 5.5%, which is close to the historical average excess return of 6.06% on the U.S. aggregate stock market index. In contrast, the equity premium on the unlevered corporate dividends is 0.62% per year in the first-best economy without limited commitment and moral hazard.

The premium on a risky asset is proportional to the covariance between the stochastic discount factor and its return. Our model generates a high equity premium for two reasons. First, agency frictions amplify the unconditional volatility of the stochastic discount factor. As explained in proposition 4, the insurance motives against persistent countercyclical tail risk in labor earnings imply a procyclical consumption share of the marginal investors. A more volatile stochastic discount factor is reflected in higher Sharpe ratios. Using the mean and the standard deviation of excess returns from table 2, the Sharpe ratio on the claim to aggregate dividends in the baseline is 48.5%, which is more than twice as large as that in the case with no frictions.

The second reason for the high equity premium is the large volatility of stock returns. In our model, stock returns are volatile because agency frictions generate fluctuations in the volatility of the stochastic discount factor over time. The general equilibrium implications of the agency problem introduce a new channel that raises the volatility of the stochastic discount factor in recessions relative to booms. The reason is the presence of the distributional state variable ϕ, whose slow-moving dynamics are summarized in persistent changes in the firm owners' share of aggregate consumption x_t. Prolonged recessions are associated with increasingly lower levels of the firm owner's consumption share. This implies that small changes in x_t translate into large variations in $\frac{x_{t+1}}{x_t} e^{g_{t+1}}$, which is the consumption growth rate of the firm owners. In equilibrium, the amplified volatility of the firm owner's consumption is compensated by a higher risk premium. The second effect of low x_t in recessions is a higher discounting of the future match surplus. This lowers firms' incentives to retain workers and exacerbates the moral hazard problem. Agents anticipate more separations and a higher downside earnings risk, which feeds back into a higher risk premia. On the other hand, in booms, the level of x_t is high, and the volatility and discounting effects are diminished.

This asymmetry results in countercyclical risk prices, high return volatility, and predictability of market returns by valuation ratios. The model delivers a 7.40% standard

[20]See Graham et al. (2015) for details on measurement of corporate leverage.

deviation of the return on the unlevered claim to corporate dividends, which is about three times higher than its counterpart in the economy with full commitment and no moral hazard. Given a low volatility of aggregate consumption and an only moderately volatile risk-free rate, most of the increase in the volatility of the market return is accounted for by the time-varying equity premium.

Table 2: AGGREGATE ASSET PRICING IMPLICATIONS

Moments	Model		Data
	Baseline	No Frictions	
Excess return on consumption			
mean	3.59%	0.62%	-
std.	7.40%	2.86%	-
Excess return on dividends			
mean	3.67%	0.62%	6.06%
std.	7.61%	2.86%	19.8%
Std of log SDF			
booms	19.15%	17.83%	38.00%
recessions	35.70%	27.80%	66.00%
Risk free rate			
mean	2.81%	5%	0.56%
std.	2.86%	0.85%	2.89%

Notes: All moments are annualized. In the "Model" column, the claim to consumption is $Y_t \int \phi_t(du)$. The the claim to dividends is $x_t Y_t$ and assumes zero financial leverage. The column labeled "No Frictions" is the first best economy, i.e., without limited commitment and moral hazard with same parameters for preferences and technology as the baseline. The column labeled "Data" column computes market return as value-weighted returns from CRSP stock index and adjusted for CPI inflation. Estimates of debt-to-equity for publicly traded U.S. firms range from 40%-50%. The risk-free rates are computed as in the appendix of Beeler and Campbell (2012). The estimates for Sharpe ratios on the market return in booms and recessions are from Lustig and Verdelhan (2012).

Time variation in the risk premium also generates the predictability of future excess returns by price-to-dividend ratios, an empirical regularity documented by several papers including Campbell and Shiller (1988), Fama and French (1988), and Hodrick (1992). In table 3, we report the results of predictability regressions in our model and those in the data. We regress excess stock market returns measured at one-to-twelve quarter horizons on the log price-to-dividend ratio at the start of the measuring period. The "Data" column reports coefficients and R^2 of these regressions using the SP500 returns over the period 1947–2015, where the data construction follows Beeler and Campbell (2012). We report the same regression results using model-simulated data in the "Model-Baseline" column. Overall, the model produces regression coefficients and R^2 that are consistent with those in the data. We also match the pattern that predictability is higher for longer-horizon returns. As a comparison, the first-best case in the column "Model-No Frictions" has a very low R^2.

Our calibration does not explicitly target moments in return predictability regressions. In fact, compared with the data, our model has a higher R^2 in predictability regressions. Leading asset pricing models, such as the habit model of Campbell and Cochrane (1999) and the long-run risk model of Bansal and Yaron (2004), typically assume that dividend growth contains an extra stochastic component that is orthogonal to shocks in aggregate consumption, which implies that a significant component stock price movements is unpredictable. For parsimony, our model does not make this assumption and therefore generates a higher predictability relative to the data.

The moments of risk-free interest rate in our model are fairly in line with standard asset pricing models. The volatility of the risk-free rate in our baseline model is 2.86% per annum. There is a wide range of estimates for this moment in the literature depending on estimation details and sample choices. The volatility of risk-free interest rate in our model is slightly higher than standard asset pricing models such as Campbell and Cochrane (1999) and Bansal and Yaron (2004). However, it is consistent with recent papers that construct "ex ante" measures of the risk-free rate; see, for example, Schorfheide et al. (2018) and Beeler and Campbell (2012).

Table 3: AGGREGATE RETURN PREDICTABILITY

Horizon	Model				Data	
(quarters)	Baseline		No Frictions			
	β	R^2	β	R^2	β	R^2
2	-0.356	0.157	-0.381	0.001	-0.062	0.042
4	-0.580	0.251	-0.739	0.001	-0.113	0.075
8	-0.788	0.329	-1.409	0.002	-0.190	0.119
12	-0.860	0.345	-2.029	0.003	-0.236	0.142
16	-0.871	0.328	-2.600	0.003	-0.277	0.166

Notes: The coefficients and R^2 of the regressions $\sum_{j=1}^{J}(r_{t+j} - r_{f,t+j}) = \alpha + \beta(pd_t) + \epsilon_{t+j}$. The column labeled "Model-Baseline" uses data simulated by the baseline calibration. The column labeled "Model-No Frictions" is the first best economy, i.e., without limited commitment and moral hazard with same parameters for preferences and technology as the baseline. The column labeled "Data" follows the construction in Beeler and Campbell (2012).

Model benchmarking In this section, we compare our results to nested cases that capture important benchmarks in the literature. The comparisons highlight why our model generates risk premia that are high and countercylical relative to setups such as those with one-sided limited commitment or exogenous wage rigidity when we require consistency with the aggregate and distributional data on labor earnings across models. Table 4 summarizes the findings.

Assume that firms can fully commit. This version of our model is similar to Alvarez and Jermann (2001) or Chien and Lustig (2010), who study the asset pricing implications of worker-side limited commitment. We keep all other features of the model unchanged, including the assumption that workers obtain all the surplus from new matches and the specification of the moral hazard problem. The results are under the column labeled "Worker-Side" in table 4.

The risk premium on the aggregate endowment claim and the volatility of returns are lower in the model with only-worker-side limited commitment. The intuition for this result can be explained as follows. First, the tightness of the worker-side limited commitment constraint does not significantly change over time. In the model, the worker-side limited commitment constraint binds for workers that receive sufficiently positive idiosyncratic shocks. However, the right tail of the distribution of idiosyncratic shocks is similar in booms and recessions. This is because our calibration is disciplined by the feature of the data that the standard deviation and the right skewness of labor earnings are almost acyclical. Second, worker-side limited commitment generates uninsurable upside risk in labor earnings. Even with recursive utility, this does not produce quantitatively significant effects on marginal utilities.

In terms of the labor market moments, we find that the model with only-worker-side limited commitment misses the large negative Kelly skewness of labor earnings in recessions and other measures of tail risk, which in our baseline model is generated by the firm-side limited commitment constraint. In addition, the lack of time variation in discount rates mitigates the cyclicality of separation rates through the moral hazard channel.

Next we compare our model to a version of Favilukis and Lin (2016b). Their model features a complete-markets stochastic discount factor and exogenous wage rigidity that generates countercyclical labor shares. We capture the Favilukis and Lin (2016b) mechanism in our setup by assuming that the aggregate dividend process follows $\tilde{x}(g_t)Y_t$, where $\tilde{x}(g_H) > \tilde{x}(g_L)$. We keep all other parameters of the model unchanged and discipline the choice of $\tilde{x}(g_H)$ and $\tilde{x}(g_L)$ by calibrating them to match the mean and standard deviation of labor shares of 67% and 2%, respectively, as in Favilukis and Lin (2016b). We then price the resulting $\tilde{x}(g)Y$ claim using a stochastic discount factor that is derived from a representative agent economy version of our model.

The "Exogenous Wage Rigidity" version of the model delivers a low equity premium of 0.681% and a small volatility of excess returns of 3.09%. These values are only slightly higher than those in our first-best case reported in table 2 under the column labeled "No Frictions." The volatility of aggregate labor share in the data is small and this limits the ability of models relying exclusively on operating leverage to generate high risk prices. In addition to wage rigidity, Danthine and Donaldson (2002) allow for exogenous movements in factor shares.

However, their setup uses standard log preferences and movement in factor shares cannot generate significant variations in the conditional volatility of the stochastic discount factor. As a result, their model cannot account for the coexistence of high volatility of returns and low volatility of dividends and the risks-free rate (Campbell and Shiller, 1988). In contrast to models with exogenous wage rigidity, in our setup, as we have shown above, risk premia are amplified primarily through the effect of agency frictions on the volatility of the stochastic discount factor. Under recursive preferences, the endogenously generated time-varying factor shares translate into persistent movements in the market price of risk and are key to produce return predictability and large variations in price-to-dividend ratios.

Our baseline generates a significantly higher premium. Agency frictions in our model amplify the volatility of the stochastic discount factor as well as the risk exposure of the aggregate dividend claim. For example, in table 2 under the column labeled "Model-Baseline," we see that while the risk premium on the aggregate consumption claim is 3.59%, the premium on the claim to corporate dividends is 3.67%. The small difference in these risk premia highlights that the amplification is primarily due to a more volatile stochastic discount factor and the role of the cash flow volatility channel is small.

Modeling the mixture distribution is necessary to match the extent and cyclicality of tail risk observed in labor earnings and, at the same time, deliver an approximately acyclical standard deviation of earnings growth as observed in the PSID. To highlight its importance, in table 4 under the column labeled "No Mixture," we report two calibrations without assuming a mixture distribution: (i) $\sigma_H = \sigma_L$ and (ii) $\sigma_H < \sigma_L$.

In the case where the distribution of idiosyncratic risk is independent of the aggregate state—that is, $\sigma_H = \sigma_L = 8.3\%$—we find that the asset pricing implications are almost similar to the first-best case, consistent with the Krueger and Lustig (2010) intuition outlined in section 4.1. In the case $\sigma_L > \sigma_H$, it is possible to make σ_L sufficiently higher than σ_H so that the implied volatility of the stochastic discount factor is similar to the baseline calibration. With $\sigma_L = 10.3\%$ and $\sigma_H = 8\%$, we are able to get an equity premium on the unlevered aggregate consumption claim of 3.2%. However, we find that the earnings growth distribution has (counterfactually) countercylical standard deviation, 38% in recessions and 30% in booms, and almost no cyclicality in Kelly skewness.

Discount rates in general equilibrium models can be constructed from the marginal rate of substitution of marginal investors. Although all agents who do not face a binding limited commitment constraint are marginal investors in our model, it is more convenient to construct the SDF from firm owners' consumption, because they are never constrained.[21] To further

[21] Firm owners are not the only investors whose marginal rate of substitution is a valid stochastic discount factor. This is true for all unconstrained workers. In our baseline calibration, less than 5% of workers are constrained in any given quarter.

illustrate the mechanism for the high volatility of SDF in our model, we report the volatility of firm owner's consumption in all four versions of the model in table 4.

First, the standard deviation of consumption growth for firm owners is 10% per year. In the data, it is difficult to reliably measure the consumption of wealthy stockholders. Using the sample from Consumer Expenditure Survey (CEX), Wachter and Yogo (2007) report that the median standard deviation of consumption growth for the bottom 50% of stock-holding households is 7.26% and that of the top 25% is about 11.38% per year. Firm owners' consumption is more procyclical than aggregate consumption. As we explain in section 4, this extra risk exposure in not insured away in equilibrium, because workers' marginal utility is high in recessions owning to uninsurable idiosyncratic risk, and it is optimal for firm owners to bear more aggregate risks than workers.

Second, the pattern of the volatility of consumption growth for firm owners echoes the pattern of the volatility of stochastic discount factor across models. The inability for other versions of the model to generate a high volatility of SDF can be attributed to the lack of risk exposure of firm owners' equilibrium consumption. The version of our model with countercyclical second moment ($\sigma_L > \sigma_H$) also generates a high volatility of consumption growth for firm owners and a high volatility of SDF, but at the expense of excessive countercyclical standard deviation in labor earnings growth.

6.2 Cross Section of Expected Returns

Value premium Stocks with low valuation ratios (value stocks) earn higher average returns than stocks with high valuation ratios (growth stocks). The difference in the mean returns of value and growth stocks is robust to various ways of constructing the valuation ratio—for example, as the ratio of the market value of the firm to its book value or as the ratio of the market price of the stock to earnings per share; see Fama and French (1992) and Fama and French (1993).

Our model generates a value premium. The price-to-earnings ratio and expected returns are functions of the state variable u_t, which summarizes the fraction of future cash flows that is promised to workers. Firms with high-u workers have a high operating leverage and a low valuation ratio. Proposition 5 states that such firms should have a higher expected return. To compare our model implications with data, we sort stocks into three portfolios ranked by earnings-to-price ratios.[22] The mean high-minus-low return is 6.27% per year, with a t-statistic of 5.01. The same portfolio-sorting procedure in the data simulated from the model generates a value premium of 4.66% per year.

[22]The return series for these portfolios is obtained from Kenneth French's website and covers the period 1956—2016.

Table 4: COMPARISON TO OTHER BENCHMARKS

Moments	Data	Models				
		Baseline	Worker-Side	Exogenous Wage Rigidity	No Mixture	
					$\sigma_H = \sigma_L$	$\sigma_H < \sigma_L$
Excess return on consumption						
mean	-	3.59%	1.16%	0.62%	1.03%	3.20%
std.	-	7.40%	2.43%	2.86%	3.16%	6.75%
Excess return on dividends						
mean	6.06%	3.67%	0.91%	0.68%	1.04%	3.27%
std.	19.8%	7.61%	2.52%	3.09%	3.36%	6.94%
Std of log SDF						
booms	38.00%	19.34%	13.75%	17.83%	9.43%	18.95%
recessions	66.00%	35.7%	23.00%	27.80%	21.22%	36.26%
Firm owners consumption						
std.	7% to 11%	10%	4.54%	3.84%	3.82%	7.72%
Risk free rate						
mean	0.56%	2.81%	5.07%	4.73%	4.33%	3.21%
std.	2.89%	2.86%	1.44%	0.39%	0.77%	2.28%

Notes: All moments are annualized. The column labeled "Data" column computes market return as value-weighted returns from CRSP stock index and adjusted for CPI inflation. Estimates of debt-to-equity for publicly traded U.S. firms range from 40%-50%. For the firm owner consumption growth in the "Data" column, we use Wachter and Yogo (2007) estimates of standard deviation of consumption growth for the stock-holding households. In the "Model" columns, the claim to consumption is $Y_t \int \phi_t(du)$. The the claim to dividends is $x_t Y_t$ and assumes zero financial leverage. For all cases, technology and preferences parameters are the same as the baseline. The column labeled "Worker-Side" relaxes constraint $v\left(u, S\right) \geq 0$. The column labeled "Exogenous Wage Rigidity" uses the first-best stochastic discount factor, in the row "Excess returns on $x_t Y_t$" we price an unlevered claim to corporate dividends with the cash flow on this claim is modeled as $\tilde{x}(g)Y$ with more details in the main text. In the column labeled "No Mixture", we set the mixture probability of drawing from the negative exponential ρ to zero. For subcolumn labeled "$\sigma_H = \sigma_L$" the value for std. in booms and recessions is set to 8.3% and for the subcolum labeled of "$\sigma_H < \sigma_L$", the value for $\sigma_L = 10.3\%$ and the value of $\sigma_H = 8.3\%$.

In our model, firms with a history of negative idiosyncratic shocks have higher expected return. A similar effect is documented by Bondt and Thaler (1985) as "long-term reversal." In our model, long-term reversal and value premium are due to the same economic mechanism, and hence they are highly correlated. Consistent with this implication of our theory, Fama and French (1996) show that the returns on value-growth portfolios and long-term reversal sorted portfolios are highly correlated.

Labor leverage and the cross section of expected returns A more direct test of the model mechanism is the connection between the value premium and firm-level obligations to workers. We use the merged CRSP/Compustat panel to test this implication.

We focus on publicly traded firms in the Compustat database and regress excess returns on a firm's equity, which are defined as the difference between equity returns and the three-month T-bill rate, on firm-level labor shares and time fixed effects.

$$\text{Excess Return}_{f,t+1} = \alpha_r + \beta_r \times \text{LaborShare}_{f,t} + \lambda_{rt}. \tag{33}$$

Following Donangelo et al. (2016), labor share for firm f at period t is constructed using

$$\text{LaborShare}_{f,t} = \frac{XLR_{ft}}{OPID_{f,t} + XLR_{f,t} + \Delta INV_{f,t}}, \tag{34}$$

where XLR is the total wage bill, OPID is operating profit before interest and depreciation, and INV is change in inventories. Whenever XLR is not available, we construct an extended labor share (ELS) using the procedure described in Donangelo et al. (2016). In table 5, we report our results both with labor share, under the column labeled "Using LS," and with extended labor share, under the column labeled "Using ELS." Consistent with our model, labor share predicts expected returns, and the point estimate for β_r is positive and significant.[23] These findings are consistent with and complementary to other studies such as Donangelo et al. (2016), who document returns on labor share–sorted portfolios and estimate versions of (33), as well as Favilukis and Lin (2016a), who use wage rigidity as a proxy for labor leverage at the industry level and show that labor leverage predicts cross-industry expected returns.

[23] The estimates are robust to including various control variables such as leverage and total assets in the regression (33). See section C in the online appendix (not for publication).

Table 5: FIRM-LEVEL RETURNS AND LABOR SHARES

Coefficients	Using LS	Using ELS
Labor share	1.38	1.25
	(0.41)	(0.19)
Time fixed effects	Yes	Yes
no. of obs.	15170	83611
no. of entities	1645	9591

Notes: The sample consist of firm-year observations from CRSP/Compustat merged files for the years 1968-2016. In the column labeled "Using LS" we use labor share computed using (34), and in the column labeled "Using ELS" we use the procedure described in Donangelo et al. (2016) and construct "extended labor share." In both specifications, labor shares are standardized and twice lagged, and standard errors are clustered at firm level.

6.3 Labor Market Implications

In this section, we focus on the implications for aggregate and cross-sectional labor market dynamics.

Discount rates and unemployment risks The incentive compatibility condition (18) links unemployment risk to worker valuations that are influenced by discount rate variations. In our model, prolonged recessions are states with high expected returns and low present values of cash flows from workers. Because firms' retention effort is not observable, they have a lower incentive to keep workers in times of low valuations. Several papers in the recent literature emphasize the link between discount rates and unemployment; see, for example, Hall (2017), Kehoe et al. (2019) and Borovicka and Borovickova (2018). In contrast to these papers, the variation in discount rates in our setting is driven by general equilibrium implications of contracting frictions, and our model is consistent with broad patterns in aggregate and cross-sectional asset returns.

In our model, average separation rates are countercyclical: 3% per year in recessions and 2% per year in booms. In the presence of separations, part of the tail risk in labor earnings is driven partly by the extensive margin when workers transition from employment to long-term unemployment. We decompose large earnings drops—that is, reductions in individual earnings of more than 20%—into two categories: separations and within-employment compensation cuts. In our calibration, 48.5% of large earnings drops are due to separation and the remaining 51.5% are due to a binding firm-side limited commitment constraint. This pattern is consistent with Guvenen et al. (2014), who document that workers in the left tail of the income distribution are more likely to experience a large drop in earnings, and claim that a nonnegligible fraction of the drop is due to unemployment risk.

The separation risk is also quantitatively important in accounting for the volatility of the stochastic discount factor. Shutting down separations—that is, under a calibration with $\theta = 1$—and keeping all other parameters unchanged lowers the annualized risk premium on the aggregate consumption claim from 3.6% to 2.14%. We interpret this as both channels being salient and quantitatively relevant in accounting for the equity premium.

Exposures to idiosyncratic and aggregate shocks Propositions 4 and 5 have direct implications for how idiosyncratic and aggregate shocks are insured in the presence of agency frictions. Owing to to firm-side limited commitment, workers with adverse histories are more exposed to idiosyncratic shocks in recessions. The optimal contract compensates this lack of insurance by providing such workers an additional hedge against aggregate shocks. Thus, the consumption of workers with adverse histories would have a relatively higher exposure to idiosyncratic shocks and a lower exposure to aggregate shocks.

To test whether firms with larger obligations to workers provide less insurance against idiosyncratic shocks, we measure the pass-through of firm-level shocks to their wage payments and check whether these pass-throughs systematically vary with the firm-level labor share. We estimate the regression

$$\begin{aligned}\Delta \log \text{WageBill}_{f,t+1} = \alpha_w + \beta_{w0}\ \text{LaborShare}_{f,t} + \beta_{w1} \Delta \log \text{Sales}_{f,t} \\ + \gamma_w \Delta \log \text{Sales}_{f,t} \times\ \text{LaborShare}_{f,t} + \lambda_{wt},\end{aligned} \tag{35}$$

where $\text{WageBill}_{f,t+1}$ is the total wage bill of firm f in year $t{+}1$ and $\text{LaborShare}_{f,t}$ is as defined in equation (34). Our sample includes all firms in Compustat for the period 1959-2017.

We report our regression results in table 6, where standard errors are in parentheses. Consistent with our model's implication of imperfect risk sharing, the point estimate of the pass-through coefficient β_1 is positive but less than one.[24] Furthermore, the interaction term $\gamma_w > 0$ and is statistically significant. This confirms the conclusion of proposition 5 that firms with higher labor leverage have a higher pass-through coefficient. In section C of the online appendix (not for publication), we estimate a version of (35) where we split the sales growth into a negative sales growth part and positive sales growth part. We find that consistent with the model, the interaction term is driven mainly by the negative part of sales growth.

[24] Guiso et al. (2005) also estimate the extent of insurance within the firm using administrative-level matched employer-employee data and similar regressions.

Table 6: FIRM-LEVEL WAGE PASS-THROUGHS AND LABOR SHARES

Coefficients	Using LS	Using ELS
LogSales	0.4159	0.3187
	(0.0422)	(0.0276)
LaborShare	-0.0726	-0.1648
	(0.007)	(0.0061)
LaborShare × LogSales	0.3871	0.3538
	(0.0776)	(0.0517)
Time fixed effects	Yes	Yes
no. of obs	40289	117128
no. of entities	4028	12806

Notes: The sample consist of firm-year observations from Compustat for the years 1959-2016. We follow Donangelo et al. (2016) in the construction of firm labor share, the results of which are reported in the column labeled "Using LS", and the construction of extended labor share, the results of which are reported in the column labeled "Using ELS." In both specifications, labor shares are twice lagged, and standard errors are clustered at the firm level.

7 Conclusion

We present an asset pricing model where risk premia are amplified by agency frictions. Under the optimal contract, sufficiently adverse shocks to worker productivity are uninsured. In general equilibrium, exposure to downside tail risk results in a more volatile stochastic discount factor and time variations in discount rates. These features of the pricing kernel yield quantitatively large and volatile risk premia and generate a substantial cross-sectional variation in returns across firms. Our model is also consistent with firm-level measures of labor share that predict both future returns and pass-throughs of firm-level shocks to wage payments.

An interesting extension of our setup would be to allow for a storage technology, such as physical capital, along with related agency frictions such as hidden savings for workers or the ability of firm owners to additionally use capital as collateral. This will open up a host of new predictions about aggregate business cycle fluctuations, firm-level and aggregate asset prices as well as capital misallocation in the cross section of firms. A recent paper by Tong and Ying (2020) builds on our setup and studies asset pricing implications of limited commitment in production economies.

References

Abraham, A., F. Alvarez-Parra, and S. Forstner (2017): "The Effects of Moral Hazard on Wage Inequality in a Frictional Labor Market," *Working Paper*.

Ai, H. and R. Bansal (2018): "Risk Preferences and the Macroeconomic Announcement Premium," *Econometrica*, 86, 1383–1430.

Ai, H. and D. Kiku (2013): "Growth to Value: Option Exercise and the Cross Section of Equity Returns," *Journal of Financial Economics*, 107, 325–349.

Ai, H., D. Kiku, R. Li, and J. Tong (forthcoming): "A Unified Model of Firm Dynamics with Limited Commitment and Assortative Matching," *The Journal of Finance*.

Ai, H. and R. Li (2015): "Investment and CEO Compensation under Limited Commitment," *Journal of Financial Economics.*, 116, 452–472.

Alvarez, F. and U. J. Jermann (2000): "Efficiency, Equilibrium, and Asset Pricing with Risk of Default," *Econometrica*, 68, 775–797.

——— (2001): "Quantitative Asset Pricing Implications of Endogenous Solvency Constraints," *Review of Financial Studies*, 14, 1117–1151.

Atkeson, A. and R. E. Lucas, Jr. (1992): "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59, 427–453.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Basak, S. and D. Cuoco (1998): "An Equilibrium Model with Restricted Stock Market Participation," *The Review of Financial Studies*, 11, 309–341.

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Berk, J. B. and J. Walden (2013): "Limited Capital Market Participation and Human Capital Risk," *Review of Asset Pricing Studies*, 3, 1–37.

Bolton, P., N. Wang, and J. Yang (2019): "Optimal Contracting, Corporate Finance, and Valuation with Inalienable Human Capital," *The Journal of Finance*, 71, 1363–1429.

Bondt, W. F. M. D. and R. Thaler (1985): "Does the Stock Market Overreact?" *The Journal of Finance*, 40, 793–805.

Borovicka, J. and K. Borovickova (2018): "Risk Premia and Unemployment Fluctuations," *Working Paper*.

Campbell, J. Y. and J. H. Cochrane (1999): "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Campbell, J. Y. and R. J. Shiller (1988): "The Dividend-price Ratio and Expectations of Future Dividends and Discount Factors," *Review of financial studies*, 1, 195–228.

Carroll, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91, 312–320.

Chien, Y. and H. Lustig (2010): "The Market Price of Aggregate Risk and the Wealth Distribution," *Review of Financial Studies*, 23, 1596–1650.

Constantinides, G. M. and D. Duffie (1996): "Asset Pricing with Heterogenous Consumers," *Journal of Political Economy*, 104, 219–240.

Constantinides, G. M. and A. Ghosh (2014): "Asset Pricing with Countercyclical Household Consumption Risk," *Working Paper*.

Danthine, J.-P. and J. B. Donaldson (2002): "Labour Relations and Asset Returns," *Review of Economic Studies*, 69, 41–64.

Davis, S. J. and T. von Wachter (2011): "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, 2, 1–72.

Donangelo, A., F. Gourio, M. Kehrig, and M. Palacios (2016): "The Cross-Section of Labor Leverage and Equity Returns," *Working Paper*, unpublished Paper.

Fama, E. and K. French (1996): "Multifactor Explanations of Asset Pricing Anomalies," *The Journal of Finance*, 51, 55–84.

Fama, E. F. and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

——— (1992): "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47, 427–465.

——— (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

Favilukis, J. and X. Lin (2016a): "Does Wage Rigidity Make Firms Riskier? Evidence from Long-horizon Return Predictability," *Journal of Monetary Economics*, 78, 80–95.

——— (2016b): "Wage Rigidity: A Quantitative Solution to Several Asset Pricing Puzzles," *Review of Financial Studies*, 29, 148–192.

Garleanu, N., L. Kogan, and S. Panageas (2012): "Displacement risk and asset returns," *Journal of Financial Economics*, 105, 491–510.

Garleanu, N. and S. Panageas (2015): "Young, Old, Conservarive and Bold. the Implications of Finite Lives and Heterogeneity for Asset Pricing," *Journal of Political Economy*, 123, 670–685.

Gomez, M. (2019): "Asset prices and wealth inequality," *Working Paper*.

Graham, J. R., M. T. Leary, and M. R. Roberts (2015): "A Century of Capital Structure: The Leveraging of Corporate America," *Journal of Financial Economics*, 118, 658–683, nBER Symposium on New perspectives on corporate capital structures.

Greenwald, D. L., M. Lettau, and S. C. Ludvigson (2016): "The Origins of Stock Market Fluctuations," *Working Paper*.

——— (2020): "How the Wealth Was Won: Factor Shares as Market Fundamentals," *Working Paper*.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the Firm," *Journal of Political Economy*, 113, 1054–1087.

Guvenen, F. (2009): "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica*, 77, 1711–1750.

Guvenen, F., S. Ozkan, and J. Song (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660.

Hagedorn, M. and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98, 1692–1706.

Hall, R. E. (2017): "High Discounts and High Unemployment," *American Economic Review*, 107, 305–330.

Harris, M. and B. Holmstrom (1982): "A Theory of Wage Dynamics," *Review of Economic Studies*, 49, 315–333.

Heaton, J. and D. J. Lucas (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, 104, 443–487.

Hodrick, R. J. (1992): "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 357–386.

Kehoe, P. J., E. Pastorino, and V. Midrigan (2019): "Debt Constraints and Employment," *Journal of Political Economy*, 127, 1926–1991.

Kehoe, T. J. and D. K. Levine (1993): "Debt-Constrained Asset Markets," *Review of Economic Studies*, 60, 865–888.

Kocherlakota, N. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63, 595–609.

Krueger, D. and H. Lustig (2010): "When Is Market Incompleteness Irrelevant for the Price of Aggregate Risk (and When Is It Not)?" *Journal of Economic Theory*, 145, 1–41.

Krusell, P. and A. Smith, Jr. (1998): "Income and Wealth Heterogenity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Lamadon, T. (2016): "Productivity Shocks, Long-term Contracts and Earnings Dynamics," *Working Paper*.

Lettau, M., S. C. Ludvigson, and S. Ma (2014): "Capital Share Risk in U.S. Asset Pricing," *NBER Working Paper*.

Lustig, H. and A. Verdelhan (2012): "Business Cycle Variation in the Risk-Return Trade-Off," *Journal of Monetary Economics*, 59, S35 – S49.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jorgensen (2009): "Long-Run Stockholder Consumption Risk and Asset Returns," *The Journal of Finance*, 64, 2427–2479.

Mankiw, N. G. (1986): "The Equity Premium and The Concentration of Aggregate Shocks," *Journal of Financial Economics*, 17, 211–219.

Mulligan, C. B. (2012): "Do Welfare Policies Matter for Labor Market Aggregates? Quantifying Safety Net Work Incentives Since 2007," *NBER Working Paper*.

Panageas, S. (2019): "The Implications of Heterogeneity and Inequality for Asset Pricing," *Working Paper*.

Rampini, A. and S. Viswanathan (2010): "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance*, 65, 2293–2322.

——— (2013): "Collateral and Capital Structure," *Journal of Finance*.

——— (2017): "Household Risk Management," *NBER Working paper 22293*.

Rudanko, L. (2011): "Aggregate and Idiosyncratic Risk in a Frictional Labor Market," *American Economic Review*, 101, 2823–2843.

Santos, T. and P. Veronesi (2020): "Leverage," *Working Paper*.

Schmidt, L. (2015): "Climbing and Falling off the Ladder : Asset Pricing Implications of Labor Market Event Risk," *Working Paper*.

Schorfheide, F., D. Song, and A. Yaron (2018): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," *Econometrica*, 86, 617–654.

Shibata, I. (2015): "Labor Market Dynamics: A Hidden Markov Approach," *Working Paper*.

Shimer, R. (2008): "The Probability of Finding a Job," 98, 268–273.

Song, J., D. Price J, F. Guvenen, N. Bloom, and T. von Wachter (2015): "Firming up Inequality," *NBER Working Paper*.

Stokey, N. L., R. E. Lucas, Jr., and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press.

Storesletten, K., C. I. Telmer, and A. Yaron (2007): "Asset Pricing with Idiosyncratic Risk and Overlapping Generations," *Review of Economic Dynamics*, 10, 519–548.

Thomas, J. and T. Worrall (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies*, 55, 541–553.

Tong, J. and C. Ying (2020): "A Dynamic-Agency Based Asset Pricing Theory with Production," *Working Paper*.

Tsuyuhara, K. (2016): "Dynamic Contracts with Worker Mobility via Directed On-the-Job Search," *International Economic Review*, 57, 1405–1424.

Veronesi, P. (2019): "Heterogeneous Households under Uncertainty," *Working Paper*.

Wachter, J. A. and M. Yogo (2007): "Why Do Household Portfolio Shares Rise in Wealth?" *Working Paper*.