Asset Pricing with Endogenously Uninsurable Tail Risk

August 26, 2019

This paper studies asset pricing in a setting in which idiosyncratic risk in human capital is not fully insurable. Firms use long-term contracts to provide insurance to workers, but neither side can commit to these contracts; furthermore, worker-firm relationships have endogenous durations owing to costly and unobservable effort. Uninsured tail risk in labor earnings arises as a part of an optimal risk-sharing scheme. In the general equilibrium, exposure to the resulting tail risk generates higher risk premia, more volatile returns, and variations in expected returns across firms. Model outcomes are consistent with the cyclicality of factor shares in the aggregate, and the heterogeneity in exposures to idiosyncratic and aggregate shocks in the cross section.

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

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 incomplete-markets-based asset pricing model that uses limited commitment and moral hazard as microfoundations to address these patterns in risk premia. Uninsured *tail* or downside risk in labor earnings arises as an outcome of optimal risk-sharing arrangements. Time variation in that tail risk drives aggregate risk prices and cross-sectional risk exposures. The model is also consistent with the cyclicality of factor shares in the aggregate, and the heterogeneity in exposures to idiosyncratic and aggregate shocks in the cross section. 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: capital owners and workers. Capital owners are well diversified and use long-term compensation contracts to provide insurance to workers against idiosyncratic fluctuations in their human capital. Two agency frictions distinguish our paper from standard representative agent asset pricing models. First, neither firm owners or workers can commit to contracts that yield continuation values lower than their outside options. Second, worker-firm relationships have endogenous durations owing to costly and unobservable effort. We embed these contracting frictions in a general equilibrium setting with aggregate shocks and then study the resulting labor market and asset pricing implications.

While worker and firm limited commitment constraints are required to match earning dynamics, downside risk in labor earnings, a key feature in the data, is driven mainly by the firm-side limited commitment and moral hazard. Compensation contracts providing perfect risk sharing would insure workers against idiosyncratic labor productivity shocks. But when firms cannot commit to negative net present value projects, large drops in labor productivity are accompanied by reductions in worker earnings. Additionally, the moral hazard problem links a firm's retention effort to the present discounted value of cash flows it expects from a worker. In periods during which future values are low, because of either low human capital or high discount rates, firms exert low effort and workers suffer higher separation risk and loss of earnings from human capital depreciation.

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. With recursive utility and persistent countercyclical idiosyncratic risks, the prospect of a future lack of risk sharing raises workers' current marginal utilities. The optimal risk-sharing scheme compensates this by allocating a higher share of aggregate output from capital 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 the capital owners. This amplifies risk prices. In our quantitative analysis, we find that Sharpe ratios are more than doubled owing to agency frictions.

Since some tail risk comes from separations, there is a feedback between limited commitment and moral hazard. The moral hazard problem links firms' retention efforts to the valuation of future cash flows they expect from workers. Higher expected returns during recessions lowers worker valuations and results in countercyclical separations. This feature of our model supplements a large literature – for example, Hall (2017) – which argues that discount rate variations are central in driving unemployment fluctuations. In our model, higher separations exacerbate tail risk and therefore the need for capital 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. Regressing returns on a claim to aggregate consumption on price-dividend ratios gives R-squares which 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.

Third, the above economic mechanism also results in a significant heterogeneity in the cross section of expected equity returns and sensitivities of wage payments to firm-level shocks. Under the optimal contract, labor compensation insures workers against aggregate productivity shocks and is countercyclical, making the residual capital income procyclical and more exposed to aggregate shocks. This delivers a form of operating leverage at the firm level. In particular, 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 large 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 predicts both future returns and pass-throughs of firm-level shocks to wage payments.

Lastly, the risk-sharing arrangement in our model is consistent with the cross-sectional variation in wealth exposure to aggregate shocks. By analyzing the consumption-replicating portfolio, we find that wealthy agents endogenously hold higher fractions of wealth in the stock market, while low-income workers invest more in the riskless asset. This is because workers who realize favorable productivity shocks are typically unconstrained, and therefore their marginal rate of substitution is equalized with those of well-diversified capital owners whose consumption is more exposed to aggregate shocks. These outcomes are in line with observations in the Survey of Consumer Finances.[1]

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. On the asset pricing side, Alvarez and Jermann (2001) and Chien and Lustig (2010) study the asset pricing implications of such environments. 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 labor income is perfectly insured. Our paper develops a model of two-sided lack of commitment as in Thomas and Worrall (1988) and augments it with moral hazard. 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.[2]

Our paper is related to asset pricing models with exogenously 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

[1] Recent work by Fagereng et al. (2016) uses administrative data on wealth and income from Norway to document that individuals with more uninsured labor income risk hold less risky portfolios.

[2] The firm-side limited commitment problem in our model has a similar structure to those studied in Bolton et al. (2014) and Ai and Li (2015). 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, on-the-job search, and search frictions. However, none of these papers allow for aggregate risks or study asset pricing.

growth that rules out the trading of financial assets in equilibrium. Heaton and Lucas (1996) and Storesletten et al. (2007) are among the few papers that depart from the no-trade equilibria to study risk premia in quantitative incomplete markets models.

In contrast to the above papers, we take an optimal contracting approach to microfound incomplete markets and use empirical evidence on labor earnings dynamics to restrict the choice of the parameters governing agency frictions. Our model allows the trading of a rich set financial assets. We explicitly characterize history dependence of labor earnings under the optimal contract. We show that the model is consistent with the empirical evidence on the cross-sectional variation in the exposures of earnings and wealth to idiosyncratic and aggregate shocks.

Theoretical predictions of our model are also consistent with a recent literature that emphasizes the importance of labor share dynamics in understanding the equity market. 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 helps explain equity and credit risk premia in production economies. The implication of our model that variations in labor shares can account for a large fraction of aggregate stock market variations is consistent with the evidence documented in Greenwald et al. (2014) 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. As in Krusell and Smith (1998), we approximate the forecasting problem facing agents.

The paper is organized as follows. We describe the environment – preferences, technology, and the contracting frictions – in section 2. In section 3, we discuss the optimal contract. In section 4, we derive the asset pricing implications that arise from agency frictions. 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 We consider a discrete time economy with $t = 0, 1, \ldots$. There are two groups of agents: a unit measure of capital 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 measure κ 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 and has an opportunity to match with a firm in a market where firms offer competitive long-term compensation contracts. Workers produce after being matched with a firm whose compensation contract they accept.

Production and human capital Production is organized within N firms.[3] We use i to index workers and $F_t(i) \in \{0, 1, \ldots N\}$ to indicate the firm where worker i is employed at time t, with the understanding that $F_t(i) = 0$ if the worker is unemployed. 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 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 $\{\pi(g'|g)\}_{g,g'}$.

A worker-firm match continues into the next period with probability $\theta_{i,t}$. With probability $1 - \theta_{i,t}$, the match dissolves and the worker becomes unemployed. Conditioning on survival of the match, the worker's human capital evolves according to

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

[3] We assume N is large so that a law of large numbers applies.

where the firm component $\eta_{F_t(i),t}$ is i.i.d. across firms but common to all workers in a firm, and the worker-specific shock, $\varepsilon_{i,t}$, is i.i.d. across workers. The shocks $\eta_{F_t(i),t}$ and $\varepsilon_{i,t}$ are independent of each other after conditioning on the aggregate shock g_t. We normalize $\mathbb{E}\left[e^{\varepsilon_{i,t}}|g_t\right] = 1$ and $\mathbb{E}\left[e^{\eta_{F_t(i),t}}|g_t\right] = 1$.

An employed worker can become unemployed either because the match dissolves exogenously with probability $1 - \theta_{i,t}$ or because the worker chooses to quit. The human capital of an unemployed worker follows

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

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

We define $\{\zeta_{i,t}\}_{t=0}^{\infty}$ to be the stochastic process that records the birth, death, and unemployment shocks experienced by worker i, with the convention that $\zeta_{i,t} = F_t(i)$ if worker i is employed by firm $F_t(i)$ in period t, and $\zeta_{i,t} = -1$ if worker i is not alive in period t. We use $z_{i,t} = \left(g_t, \eta_{F_t(i),t}, \varepsilon_{i,t}, \zeta_{i,t}\right)$ to denote time t shocks for a worker and $z_i^t = \left(g^t, \eta_{F_t(i)}^t, \varepsilon_i^t, \zeta_i^t\right) = \left\{g_s, \eta_{F_s(i),s}, \varepsilon_{i,s}, \zeta_{i,s}\right\}_{s=0}^{t}$ to denote the history of shocks up to time t. Because all workers are endowed with one unit of human capital at birth, given the history of shocks, z_i^t, we can recover $h_{i,t}$ for all t using equations (1) and (2), which we denote as $h_t(z_i^t)$.

Matching and separation We assume that firms affect the survival probability of a match with a worker by exerting costly effort. We denote the effort for keeping worker i at time t as $\theta_{i,t}$ and assume that the cost of effort per unit of output is specified by a function $A(\theta)$ with strictly positive first three derivatives for all $\theta \in (0, 1)$.

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. We assume that there is no cost for posting vacancies and all firms can compete for new workers.

A contract specifies both the compensation to the worker $\mathbf{C} \equiv \left\{C_t\left(z^t\right)\right\}_{t=0}^{\infty}$ and the employer's effort for keeping the worker $\boldsymbol{\theta} \equiv \left\{\theta_t\left(z^t\right)\right\}_{t=0}^{\infty}$, as functions of the history of shocks. We denote a contract using $\langle \mathbf{C}, \boldsymbol{\theta} \rangle$. In principle, we could adopt a more general contracting space that allows a worker to be paid by all N firms. However, it is straightforward to show that because of limited commitment (which we introduce next), only the employer firm will pay the worker and exert a nontrivial retention effort. Hence, without loss of generality, $\langle \mathbf{C}, \boldsymbol{\theta} \rangle$ needs only to specify the payment and retention effort between the worker and his employer as functions of the worker's history.[4]

Capital owners Capital owners are endowed with ownership of firms and have no labor income. There is a competitive market where capital owners can trade a full set of one-period ahead Arrow securities. We use $X_t(g^t)$ to denote the aggregate consumption of the capital owners and $\Lambda_t(g^t)$ to denote the stochastic process for state prices.

Contracting frictions Let $\langle \mathbf{C}, \boldsymbol{\theta} \rangle$ be a contract and let $V_t(z^t|\mathbf{C}, \boldsymbol{\theta})$ be the value of the contract to the employer of a worker with history z^t. The values $\left\{V_t(z^t|\mathbf{C}, \boldsymbol{\theta})\right\}_{t=0}^{\infty}$ satisfy the recursion

$$V_t(z^t|\mathbf{C}, \boldsymbol{\theta}) = \left[Y_t h_t(z^t)\left[1 - A\left(\theta_t\left(z^t\right)\right)\right] - C_t\left(z^t\right)\right] + \kappa\theta_t\left(z^t\right)\mathbb{E}\left[\frac{\Lambda_t(g^{t+1})}{\Lambda_t(g^t)}V_{t+1}(z^{t+1}|\mathbf{C}, \boldsymbol{\theta})\middle| z^t\right]. \tag{3}$$

Because shocks are i.i.d. across firms and workers, the value function depends only on the history and not the identity of the worker.

Let $U_t^*(z^t)$ be the maximum utility a worker can achieve in the labor market at time t at history z^t by matching with a firm. The utility of an unemployed worker, denoted $\overline{U}_t(z^t)$, can be constructed recursively according to

$$\overline{U}_t(z^t) = \left[(1-\beta)\left(bY_t h_t(z^t)\right)^{1-\frac{1}{\psi}} + \beta\overline{\mathbb{M}}_t^{1-\frac{1}{\psi}}(z^t)\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{4}$$

where $\overline{\mathbb{M}}_t(z^t) = \left(\kappa\mathbb{E}\left[(1-\chi)\overline{U}_{t+1}^{1-\gamma}(z^{t+1}) + \chi U_{t+1}^{*1-\gamma}(z^{t+1})\middle| z^t\right]\right)^{\frac{1}{1-\gamma}}$ is the certainty equivalent of the next period utility.

Given the contract $\langle \mathbf{C}, \boldsymbol{\theta} \rangle$, the utility of a worker at history z^t satisfies

$$U_t(z^t|\mathbf{C}, \boldsymbol{\theta}) = \left[(1-\beta)\,C_t\left(z^t\right)^{1-\frac{1}{\psi}} + \beta\mathbb{M}_t(z^t)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{5}$$

with $\mathbb{M}_t(z^t) = \left(\kappa\mathbb{E}\left[\theta_t\left(z^t\right)U_{t+1}^{1-\gamma}(z^{t+1}|\mathbf{C}, \boldsymbol{\theta}) + \left(1-\theta_t\left(z^t\right)\right)\overline{U}_{t+1}^{1-\gamma}\left(z^{t+1}\right)\middle| z^t\right]\right)^{\frac{1}{1-\gamma}}$.[5]

[4] To simplify notation, we assume that workers with the same history receive the same contract. This assumption does not affect the quantitative implications of our model since we focus on the steady state. Conceptually, this assumption can be easily generalized by indexing workers by their time-0 discounted utility, as in Atkeson and Lucas (1992).

[5] The death rate κ of workers does not affect the relative discount rates of firms and workers as κ appears

There are two types of agency frictions. First, neither firms nor workers can fully commit. At the beginning of each period, before production takes place, firms and workers have an opportunity to terminate relationship unilaterally and take their outside options. Second, firms' choices of effort $\{\theta_{it}\}_{i,t}$ are observable neither to workers nor to any other firms. The presence of agency frictions imposes incentive compatibility constraints on the contracts offered, which we describe next.

Upon termination of the contract, the firm can either keep the position vacant or hire a new worker. Perfect competition on the labor market and no cost for keeping or posting vacancies imply that the value of firms' outside options is zero. Thus, the firm-side limited commitment constraint on continuation values becomes[6]

$$V_t(z^t|\mathbf{C},\boldsymbol{\theta}) \geq 0 \quad \forall z^t. \tag{6}$$

If a worker chooses to separate after history z^t, he becomes unemployed and obtains utility $\overline{U}_t(z^t)$. Therefore, the worker-side limited commitment constraint becomes

$$U_t(z^t|\mathbf{C},\boldsymbol{\theta}) \geq \overline{U}_t(z^t) \quad \forall z^t. \tag{7}$$

Finally, the fact that $\boldsymbol{\theta}$ is not observable to workers implies that the choice of $\boldsymbol{\theta}$ must be incentive compatible from the firm's perspective. That is, $\forall z^t$, $\forall \tilde{\theta} \in [0,1]$[7]

$$V_t(z^t|\mathbf{C},\boldsymbol{\theta}) \geq \left[Y_t h_t(z^t)\left[1 - A\left(\tilde{\theta}\right)\right] - C_t\left(z^t\right)\right] + \kappa\tilde{\theta}\mathbb{E}\left[\frac{\Lambda_t(g^{t+1})}{\Lambda_t(g^t)} V_{t+1}(z^{t+1}|\mathbf{C},\boldsymbol{\theta})\middle| z^t\right]. \tag{8}$$

Given a pricing kernel $\left\{\Lambda_t(g^t)\right\}_t$ and maximum utilities $\left\{U_t^*(z^t)\right\}_t$ that workers can obtain on the labor market, we can construct firm values $\left\{V_t(z^t|\mathbf{C},\boldsymbol{\theta})\right\}_t$, and worker utilities $\left\{U_t(z^t|\mathbf{C},\boldsymbol{\theta})\right\}_t$, for all histories z^t under a contract $\langle\mathbf{C},\boldsymbol{\theta}\rangle$. We next define a feasible contract.

Definition 1. *A contract $\langle\mathbf{C},\boldsymbol{\theta}\rangle$ is said to be feasible with respect to $\left\{\Lambda_t(g^t)\right\}_t, \left\{U_t^*(z^t)\right\}_t$ if it satisfies limited commitment constraints (6) and (7) and incentive compatibility constraints (8), where the worker's outside option $\overline{U}_t(z^t)$ in equation (7) satisfies (4).*

in both the firm's value function in equation (3) and the worker's utility recursion in equation (5).

[6] The setting with firm-side limited commitment can also be interpreted as an environment in which firms borrow subject to an endogenously specified limit on collateral. In particular, our formulation is equivalent to one where only the NPV of the firm's cash flow can be used as collateral. For models of limited collateral, see Lustig and Van Nieuwerburgh (2005) and Rampini and Viswanathan (forthcoming). We thank an anonymous referee for pointing out this connection.

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

An equilibrium consists of state prices $\left\{\Lambda_t(g^t)\right\}_t$, the maximum utility $\left\{U_t^*(z^t)\right\}_t$ that a worker with an employment opportunity can achieve, optimal contracts $\left\langle \widehat{\mathbf{C}},\widehat{\boldsymbol{\theta}}\right\rangle = \left\{\widehat{C}_t\left(z^t\right), \widehat{\theta}_t\left(z^t\right)\right\}_{t=0}^{\infty}$ that maximize firm value among all feasible contracts, and a consumption process for the capital owners $\left\{X\left(g^t\right)\right\}_t$. Below, we define and study a recursive competitive equilibrium.

2.2 Recursive Formulation

State variables Equilibrium Arrow 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. We use homotheticity properties of preferences and technology to construct a recursive competitive equilibrium where the history of aggregates can be summarized by state variables (ϕ, g, B) and the history for an individual worker can be summarized by a single state variable u. Here, ϕ is a one-dimensional distribution of agent types, g is the Markov state of aggregate productivity, B is the total compensation to all unemployed workers normalized by aggregate productivity, and u is the current-period continuation utility normalized by human capital and aggregate productivity. That we ultimately need to keep track of only a one-dimensional distribution as a state variable is key for our quantitative analysis.

Let u_t be a worker's period t utility U_t divided by human capital and aggregate productivity $Y_t h_t$. Given a feasible contract $\langle \mathbf{C}, \boldsymbol{\theta}\rangle$, the individual state variable u_t can be constructed from z^t and denoted as $u_t = u_t\left(z^t \middle| \mathbf{C}, \boldsymbol{\theta}\right)$. The aggregate state variables ϕ_t and B_t can also be constructed recursively from the history of aggregate shocks, which we denote as $\phi_t\left(g^t\right)$ and $B_t\left(g^t\right)$. In our construction, $U_t^*(z^t)$ and $\overline{U}_t(z^t)$ take the form of

$$U_t^*(z^t) = u^*\left(\phi_t\left(g^t\right), g_t, B_t\left(g^t\right)\right) h_t(z^t) Y_t, \quad \overline{U}_t(z^t) = \overline{u}\left(\phi_t\left(g^t\right), g_t, B_t\left(g^t\right)\right) h_t(z^t) Y_t.$$

Equation (4) implies the following relationship between $\overline{u}\left(\phi, g, B\right)$ and $u^*\left(\phi, g, B\right)$:

$$\overline{u}\left(\phi, g, B\right) = \left[(1-\beta)\, b^{1-\frac{1}{\psi}} + \beta\left[\lambda \overline{m}\left(\phi, g, B\right)\right]^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{9}$$

with

$$\overline{m}\left(\phi, g, B\right) \equiv \left(\kappa \mathbb{E}\left[e^{(1-\gamma)g'}\left\{(1-\chi)\overline{u}^{1-\gamma}\left(\phi', g', B'\right)^{1-\gamma} + \chi u^{*1-\gamma}\left(\phi', g', B'\right)\right\}\middle| g\right]\right)^{\frac{1}{1-\gamma}}.$$

Workers' utility can also be represented in normalized terms as

$$U_t(z^t|\mathbf{C},\boldsymbol{\theta}) = u_t\left(z^t\middle|\,\mathbf{C},\boldsymbol{\theta}\right)h_t\left(z^t\right)Y_t.$$

Recursive optimal contracting Let $\Lambda\left(g'\middle|\,\phi,g,B\right)$ be one-period-ahead Arrow security price, $s' = (g',\eta',\varepsilon')$ the vector of the realization of next-period shocks, $\Omega(ds'|g)$ the distribution of s' given the current aggregate state g, and (g',ϕ',B') the next-period aggregate states. The normalized firm value $v\left(u|\phi,g,B\right)$ satisfies a Bellman equation

$$v\left(u|\phi,g,B\right) = \max_{c,\theta,\{u'(s')\}_{s'}} \begin{array}{c} 1-c-A(\theta)+ \\ \kappa\theta\int\Lambda\left(g'|\phi,g,B\right)e^{g'+\eta'+\varepsilon'}v\left(u'\left(s'\right)|\phi',g',B'\right)\Omega(ds'|g), \end{array} \tag{10}$$

where the maximization is subject to

$$u = \left[(1-\beta)\,c^{1-\frac{1}{\psi}} + \beta m^{1-\frac{1}{\psi}}\left(u|\,\phi,g,B\right)\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{11}$$

$$v\left(u'\left(s'\right)\middle|\,\phi',g',B'\right) \geq 0,\ for\ \ all\ \ s', \tag{12}$$

$$u'(s') \geq \lambda\overline{u}\left(\phi',g',B'\right),\ for\ \ all\ \ s', \tag{13}$$

$$A'(\theta) = \kappa\int\Lambda\left(g'|\phi,g,B\right)e^{g'+\eta'+\varepsilon'}v\left(u'\left(s'\right)|\phi',g',B'\right)\Omega(ds'|g), \tag{14}$$

and $m(u|\phi,g,B)$ in the promise-keeping constraint (11) is defined as

$$m\left(u|\,\phi,g,B\right) = \left\{\kappa\int e^{(1-\gamma)(g'+\eta'+\varepsilon')}\left[\theta\left[u'\left(s'\right)\right]^{1-\gamma} + (1-\theta)\left[\lambda\overline{u}\left(\phi',g',B'\right)\right]^{1-\gamma}\right]\Omega(ds'|g)\right\}^{\frac{1}{1-\gamma}}.$$

Inequalities (12) and (13) are the recursive counterparts of the limited commitment constraints (6) and (7). Equation (14) is the first-order necessary condition for firms' choice of retention effort. Because the cost function $A\left(\theta\right)$ is strictly convex, first-order conditions (14) are equivalent to (8) and, therefore, necessary and sufficient for incentive compatibility. We label the above maximization problem as *P1*.

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

$$w\left(\phi,g,B\right) = \left[(1-\beta)\,x\left(\phi,g,B\right)^{1-\frac{1}{\psi}} + \beta n^{1-\frac{1}{\psi}}\left(\phi,g,B\right)\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{15}$$

with the certainty equivalent $n\left(\phi,g,B\right) = \left\{\kappa\sum_{g'}\pi\left(g'|g\right)e^{(1-\gamma)g'}w^{1-\gamma}\left(\phi',g',B'\right)\right\}^{\frac{1}{1-\gamma}}$.

Finally, we describe the construction of the aggregate distributional state variable ϕ, which we will refer to as the "summary measure." Let $\Phi_j(du, dh)$ denote the joint distribution of (u, h) for workers in firm j and $\Phi_0(dh)$ the distribution of human capital of unemployed workers.[8] In general, $\{\Phi_j\}_{j=0}^N$ is a state variable in the construction of a recursive equilibrium because the resource constraint,

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

depends on $\{\Phi_j\}_{j=0}^N$. Let $c(u|\, \phi, g, B)$ attain the optimal value in the problem *P1*. The total compensation to all workers

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

where we decompose the joint distributions into a marginal distribution and a conditional distribution: $\Phi_j(du, dh) = \Phi_j(dh|\, u)\, \Phi_j(du)$. We define the summary measure by $\phi(du) \equiv \sum_{j=1}^{n} \int h\Phi_j(dh|\, u)$ 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. Total consumption equals $Y \int \phi(du)$.

We define the total compensation to all unemployed workers normalized by aggregate productivity as $B = \int bh\Phi_0(dh)$. The resource constraint can be written as

$$B + \int [c(u|\phi, g, B) + A(\theta(u|\phi, g, B))]\, \phi(du) + x(\phi, g, B) = \int \phi(du). \qquad (16)$$

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 $c(u|\, \phi, g, B)$, $\theta(u|\, \phi, g, B)$, $\{u'(u, s'|\, \phi, g, B)\}_{s'}$ by solving problem *P1*. In the second, we use the policy functions to construct the law of motion of the endogenous state variables u, ϕ, and B. The summary measure ϕ has a continuous density on $[\lambda\overline{u}(\phi, g, B),\ u^*(\phi, g, B))$ and the density of next period summary

[8] We do not need to assign a promised utility to unemployed workers because their compensation depends only on their human capital.

measure ϕ' in state g' is

$$\phi'(d\tilde{u}) = (1-\kappa)\int \theta(u|\phi,g,B)\left[\iint e^{\varepsilon'+\eta'} f(\varepsilon',\eta'|g')\, I_{\{u'(u,s'|\phi,g,B)\in d\tilde{u}\}} d\varepsilon' d\eta'\right]\phi(du), \tag{17}$$

for $\forall \tilde{u} \in [\lambda\overline{u}(\phi',g',B'),\ u^*(\phi',g',B'))$, where I is the indicator function. Entry of newly employed workers puts a mass point on $u^*(\phi,g,B)$ and $\phi'(u^*(\phi',g',B')) = \kappa + (1-\kappa)\lambda\frac{B}{b}$. The law of motion for B is given by

$$B' = \kappa\lambda\left[B(1-\chi) + b\int[1-\theta(u|\phi,g,B)]\phi(u)\,du\right]. \tag{18}$$

Definition 2. *A recursive competitive equilibrium consists of state prices* $\{\Lambda(g'|\phi,g,B)\}_{g'}$, *workers' outside option* $\overline{u}(\phi,g,B)$, *the utility* $u^*(\phi,g,B)$ *of newly employed workers, firm values* $v(u|\phi,g,B)$ *and policy functions* $c(u|\phi,g,B)$, $\theta(u|\phi,g,B)$, $\{u'(u,s'|\phi,g,B)\}_{s'}$, *consumption share of capital owners* $x(\phi,g,B)$, *and a law of motion for* (ϕ,B), *such that*

1. *The stochastic discount factor* $\Lambda(g'|\phi,g,B)$ *is consistent with capital owners' consumption:*[9]

$$\Lambda\left(g'\middle|\phi,g,B\right) = \beta\left[\frac{x(\phi',g',B')e^{g'}}{x(\phi,g,B)}\right]^{-\frac{1}{\psi}}\left[\frac{w(\phi',g',B')e^{g'}}{n(\phi,g,B)}\right]^{\frac{1}{\psi}-\gamma}, \tag{19}$$

   where capital owners' utility $w(\phi,g,B)$ *and certainty equivalent* $n(\phi,g,B)$ *are defined in equations (15).*

2. *Given* $\Lambda(g'|\phi,g,B)$, *the law of motion for* (ϕ,B), *and a worker's outside value* $\overline{u}(\phi,g,B)$, *the value function and the policy functions solve problem P1.*

3. *Given the policy functions, the law of motion for* (ϕ,B) *satisfies (17) and (18).*

4. *Functions* $\overline{u}(\phi,g,B)$ *and* $u^*(\phi,g,B)$ *satisfy (9), and for all* (ϕ,g,B),

$$\begin{aligned} u^*(\phi,g,B) &= \text{argmax}_{\tilde{u}} v(\tilde{u}|\phi,g,B) \\ s.t.\quad & v(\tilde{u}|\phi,g,B) \geq 0 \end{aligned} \tag{20}$$

5. *The policy functions, the summary measure* ϕ, *and the compensation for unemployed workers* B *satisfy the resource constraint (16).*

[9] For brevity,we specify the stochastic discount factor as a function of the capital owners' consumption directly without explicitly specifying the capital owners' consumption and portfolio problem. Because the capital owners are well diversified, their consumption and investment choices are standard.

3 The Optimal Contract

With full commitment, firms can perfectly insure workers against idiosyncratic shocks and thus assure that workers' continuation utilities do not respond to idiosyncratic shocks, that is, $e^{\varepsilon'+\eta'}u'\left(g',\eta',\varepsilon',u|\,\phi,g,B\right)$ is equalized across all possible realizations of η' and ε'. When 0 is a possible realization of $\eta'+\varepsilon'$, this optimal risk-sharing condition can be written as

$$u'\left(u,g',\eta',\varepsilon'\middle|\,\phi,g,B\right)=e^{-\varepsilon'-\eta'}u'\left(u,g',0,0\middle|\,\phi,g,B\right),\forall g',\eta',\varepsilon'. \tag{21}$$

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

Under limited commitment, equation (21) cannot hold for all η',ε'. For example, a sufficiently negative realization of η' or ε' will make the firm-side limited commitment constraint (12) bind. Perfect risk sharing means that workers consume the same fraction of aggregate consumption at all times. 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 wages.[10] The next proposition summarizes properties of the optimal contract.

Proposition 1. *Suppose that there exists an equilibrium in which the stochastic discount factor and the law of motion for aggregate state variables satisfy condition (A.1) in Appendix A1. Then there exist* $\underline{\varepsilon}(u,g'|\phi,g,B)$ *and* $\overline{\varepsilon}(u,g'|\phi,g,B)$ *such that* $\underline{\varepsilon}(u,g'|\phi,g,B)<\overline{\varepsilon}(u,g'|\phi,g,B)$ *and*

1. For all $\varepsilon'+\eta'\in(-\infty,\ \underline{\varepsilon}(u,g'|\phi,g,B))\cup(\overline{\varepsilon}(u,g'|\phi,g,B),\ \infty)$,

$$u'(u,s'|\phi,g,B)=\begin{cases}u^{*}\left(\phi',g',B'\right) & \varepsilon'+\eta'\leq\underline{\varepsilon}(u,g'|\phi,g,B),\\ \lambda\overline{u}\left(\phi',g',B'\right) & \varepsilon'+\eta'\geq\overline{\varepsilon}(u,g'|\phi,g,B).\end{cases} \tag{22}$$

2. For all $\varepsilon'+\eta'\in[\underline{\varepsilon}(u,g'|\phi,g,B),\overline{\varepsilon}(u,g'|\phi,g,B)]$, $u'(s',u|\phi,g,B)$ is strictly decreasing in $\varepsilon'+\eta'$ and satisfies

$$\begin{aligned}&\left[\frac{x\left(\phi',g',B'\right)}{x\left(\phi,g,B\right)}\right]^{-\frac{1}{\psi}}\left[\frac{w\left(\phi',g',B'\right)}{n\left(\phi,g,B\right)}\right]^{\frac{1}{\psi}-\gamma}\left(1+\frac{\iota\left(u|\,\phi,g,B\right)}{\theta\left(u|\,\phi,g,B\right)}\right)\\ =\;& e^{-\gamma(\eta'+\varepsilon')}\left[\frac{c\left(u'\left(u,s'|\,\phi,g,B\right),\phi',B'\right)}{c\left(u|\,\phi,g,B\right)}\right]^{-\frac{1}{\psi}}\left[\frac{u'\left(u,s'|\,\phi,g,B\right)}{m\left(u|\,\phi,g,B\right)}\right]^{\frac{1}{\psi}-\gamma},\end{aligned} \tag{23}$$

[10] More formally, the function $v(u|\phi,g,B)$ is bounded above by the first best $v^{FB}(u|\phi,g,B)$, which is linear in u with a slope of -1. Hence, as u approaches ∞, firm values will be negative.

where $\iota\left(u|\,\phi,g,B\right)>0$ is given in Appendix A1.

3. Firms' optimal effort $\theta(u|\phi,g,B)$ is decreasing in u.

Proof. See Appendix A1. □

The above proposition has several implications. First, extremely large and extremely small realizations of ε' and η' both lead to binding limited commitment constraints and therefore cannot be hedged. Equation (21) suggests that to provide insurance to workers, positive realizations of $\eta'+\varepsilon'$ must be offset by decreases in $u'\left(u,s'|\,\phi,g,B\right)$. The limited commitment constraint on the worker's side, $u\left(s'\right)\geq\lambda\overline{u}\left(\phi',g',B'\right)$, imposes a lower bound on $u(s')$, which means that unnormalized continuation utility must increase after extremely large realizations of $\eta'+\varepsilon'$. High promised values are met with higher future wages. This feature of our setting is similar to that in Harris and Holmstrom (1982), Kehoe and Levine (1993), and Alvarez and Jermann (2000). In contrast to these papers in which workers are perfectly insured against downside risk, the limited commitment constraint on the firm side implies that a sufficiently negative $\eta'+\varepsilon'$ such that (12) binds will result in permanent reductions in compensation.

Second, in the interior of $(\underline{\varepsilon}(u,g'|\phi,g),\overline{\varepsilon}(u,g'|\phi,g))$, the intertemporal marginal rate of substitution of a worker does not depend on idiosyncratic shocks η' and ε'. This resembles the perfect risk-sharing condition (21). Because the consumption policy $c\left(u|\,\phi,g,B\right)$ is strictly increasing in u, the optimal risk-sharing condition (23) implies that normalized continuation utilities $u'\left(u,s'|\,\phi,g,B\right)$ are strictly decreasing in $\eta'+\varepsilon'$. As a result, the promised value u for a worker-firm pair that realizes a history of negative productivity shocks will drift upwards.

Incentive compatibility constraint (14) requires that the marginal cost $A'(\theta)$ of retaining the worker equals its marginal benefit, the present value of the cash flow that the worker can bring to the firm, $\kappa\int\Lambda\left(g'|\phi,g,B\right)e^{g'+\eta'+\varepsilon'}v\left(u'\left(s'\right)|\phi',g',B'\right)\Omega(ds'|g)$. Firm effort θ is smaller than its first-best counterpart because the social benefit also include workers' utility gain by staying employed. The optimal contract manages this trade-off by back-loading firms' dividend payouts and front-loading workers' consumptions relative to the first-best case. Back-loading introduces a wedge $\left(1+\frac{\iota(u|\phi,g,B)}{\theta(u|\phi,g,B)}\right)$ between marginal rate of substitutions of the capital owner and workers, where the term $\iota\left(u|\,\phi,g,B\right)$ is the Lagrangian multiplier on constraint (14).

Finally, part 3 of Proposition 1 implies that separation rates are higher for unproductive worker-firm matches. Workers who experienced a sequence of negative productivity shocks have low human capital, a high u, and a lower future surplus $v(u'(s'))$ for the firm. It is less profitable for firms to keep such workers. Incentive constraint (14) implies that the optimal choice of θ must be low. More generally, separation rates are higher when the value of the worker to the firm is lower. This may be due to either a lower future surplus from the worker (that is, lower levels of $v(u'(s'))$) or a higher discount rate (that is, lower values of Λ).

4 Agency Frictions and Asset Pricing

In this section, we highlight how limited commitment and moral hazard 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 risks 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 to distinguish it from alternatives in the literature. We also derive a set of testable predictions of our model mechanism which we later confront 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 $f(\varepsilon, \eta| g)$, is independent of 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 to our setting with contracting frictions, we start with a definition.

Definition 3. *An equivalent deterministic economy with a modified discount rate is the economy described in section 2.2 with no aggregate growth and a modified discount rate* $\hat{\beta} = \beta \left(\mathbb{E}\left[e^{(1-\gamma)g'} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}$*, with* $\mathbb{E}$ *being the unconditional expectations operator.*

In the following proposition we show that equilibrium allocations and state prices in the stochastic economy can be constructed from the equilibrium of an equivalent deterministic economy with a modified discount rate.

Proposition 2. *(Krueger and Lustig) Suppose that* g_t *is i.i.d. over time and that* $f(\varepsilon, \eta| g)$ *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 stochastic economy described in section 2.2 with the stochastic discount factor satisfying*

$$\Lambda\left(g'\middle|\,\phi, g, B\right) = \frac{1}{\hat{R}\left(\phi, B\right)} \frac{e^{-\gamma g'}}{\mathbb{E}\left[e^{(1-\gamma)g'}\right]}, \tag{24}$$

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

Proof. See Appendix A1. □

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 $\beta e^{-\gamma g'}$. This is also the stochastic discount factor for the representative agent economy in which the growth rate of aggregate consumption is g_t. Equation (24) 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. We show in Appendix A1 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 capital owners in the stochastic economy equals that in the equivalent deterministic economy.

4.2 Aggregate Implications

Proposition 2 tells us that to understand the impact of agency frictions on aggregate risk premia, we must deviate from its assumptions of i.i.d. growth and the time-invariant distribution of idiosyncratic shocks. In the rest of this section, we analyze a special case of our model that highlights the interaction between agency frictions, labor earnings, and the market price of aggregate risks. We proceed by making several simplifying assumptions. These assumptions are designed to isolate features and implications that are novel to our setting, and to help us 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 $\pi\left(g'\middle|\,g\right) = 1$ 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.*[11]

This assumption includes the main departures from Proposition 2. 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. The assumption that firm-level shocks $\eta = 0$ is without loss of generality, since Proposition 1 shows that the optimal contract depends only on $\varepsilon + \eta$. In what follows, we interpret ε as both a firm-level shock 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 1, uninsurable risk in the left tail of labor earnings comes from the firm-side limited commitment. In section 6.1, we show the that worker-side limited commitment has little 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 the worker side.

Assumption 4. *Effort is only costly 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, there will be no risk sharing and workers consume a fixed fractions of their outputs. This assumption captures that workers' consumption is more exposed to idiosyncratic shocks in future recessions because of lack of risk sharing. 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 from period 2 on.

[11] See Appendix A2 for the definition and the properties of the negative exponential distribution.

We plot an event tree for the simple economy in figure 1. Let capital owners' consumption share at date 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 clears the market. 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 $x_L = x(g_L)$ denote the capital owners' consumption share at nodes H and L, respectively. 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, $c_H(u_0) \equiv c(u'(u_0, g_H)|\, g_H)$ and $c_L(u_0, \varepsilon) \equiv c(u'(u_0, g_L, \varepsilon)|\, g_L)$ to denote the compensation policy, and $v_H(u_0) \equiv v(u'(u_0, g_H)|\, g_H)$ and $v_L(u_0, \varepsilon) \equiv v(u'(u_0, g_L, \varepsilon)|\, g_L)$ to denote firms' value function at nodes H and L, respectively. 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 3. *(Aggregate Implications) Under Assumptions 1-5, for expected utility preferences, i.e., $\gamma = 1$, capital owners' consumption share is countercyclical, that is, $x_H < x_L$. For general recursive utility with $\gamma \geq 1$, there exists a $\hat{\gamma} \in [1, 1 + \xi)$ such that if $\gamma > \hat{\gamma}$, then (i) capital owners' consumption share is procyclical, that is, $x_H > x_L$ and (ii) separation rates are countercyclcal, 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 capital owners, amplification in the market price of risk relative to a representative agent model is equivalent to capital owner's consumption share being procyclical. The first part of Proposition 3 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 1). 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 a larger fraction of agents get constrained in recessions relative to booms. Because constrained firms cut compensation, in the aggregate there are more resources available. Since goods markets need to clear, these resources are allocated between the capital 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 capital owners and the unconstrained agents. Therefore, for $\gamma = 1 = \frac{1}{\psi}$, the consumption share of both capital owners and unconstrained agents must increase and $x_L > x_H$.

The second implication of the Proposition 3 is that keeping the intertemporal elasticity of substitution fixed, a large enough risk aversion results in a procyclical consumption share for capital owners. As risk aversion exceeds the inverse of the intertemporal elasticity of substitution, contemporaneous marginal utilities are decreasing functions of continuation utility. This forward looking property of the preferences is what translates uninsurable tail risk in labor earnings into a higher market price of aggregate shocks.[12]

Persistent recessions that are associated with a lack of risk sharing in the future imply lower continuation values and higher marginal utilities in the current period for workers. Optimal risk sharing which requires equating marginal rates of substitution between capital owners and unconstrained workers is now achieved by transferring resources away from the capital owners. Proposition 3 says that for sufficiently high risk aversion, this incentive is strong enough to dominate the effect of market clearing and delivers a procylical consumption shares for capital owners.

The last part of Proposition 3 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 1, under firm-side limited commitment, this tail risk cannot be fully insured within optimal labor compensation contracts. Second, workers become unemployed with probability θ in each period. The countercyclicality of unemployment risk asserted in part (ii) of Proposition 3 is a direct consequence of incentive compatibility under moral hazard. Equation (14) requires firms to equalize the marginal cost of retention effort to its marginal benefit. The marginal benefit of retention is the present value of profits that a worker can create for the firm. Valuation ratios in recessions are lower relative to booms. Thus, firms have less incentive to exert costly effort to retain workers in recessions relative to booms leading to countercyclical separation rates.

The effects of limited commitment and that 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 capital 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

[12] 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 $\gamma > 1$ that is responsible for the procyclical consumption share in our model.

more separations.

Contrasting the mechanism to alternatives proposed in the literature The above result is in contrast with several exogenously incomplete market models, for example, Constantinides and Duffie (1996), Constantinides and Ghosh (2014), and Schmidt (2015). In those papers, all agents are marginal investors in risky assets, and hence countercyclical uninsurable risk in consumption automatically translates into a more volatile pricing kernel. In the simple example where market incompleteness is determined by optimal contracting under agency frictions, agents with adverse idiosyncratic shocks are constrained and not marginal. Hence, higher idiosyncratic volatility by itself is not sufficient to 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 constraint. 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 binds for worker-firm pairs that receive large positive idiosyncratic productivity shocks. Constrained workers need to be compensated with higher current and future wages. This lowers the consumption for unconstrained agents, 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 in amplifying the volatility of the stochastic discount factor than upside risk. 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 3 also distinguishes our model from Danthine and Donaldson (2002), 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 premium 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 only affects 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.[13] 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 formally test them using panel data in section 6.3.

In our model, heterogeneity in firms is summarized by a single state variable u. High-u firms promise a larger fraction of cash flow to workers than low-u firms. Thus u can be interpreted as "labor leverage". Below we provide two comparative static results with respect to u_0.

Proposition 4. *(Cross-Sectional Implications) Let $c_L(u_0,\varepsilon) \equiv c\left(u'(u_0, g_L, \varepsilon) | g_L\right)$ denote the compensation policy for workers with initial promised utility u_0 at time 0. Under Assumptions 1-5, (i) the elasticity of wage payments with respect to idiosyncratic shocks*

$$\frac{\partial}{\partial u_0}\mathbb{E}\left[\frac{\partial \ln\left[e^{\varepsilon} c_L(u_0,\varepsilon)\right]}{\partial \varepsilon}\right] > 0. \tag{25}$$

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 as $\underline{\varepsilon}(\hat{u}, g_L) = \ln\frac{1+\xi}{\xi}$, such that $\forall u_0 < \hat{u}$,

$$\frac{\partial}{\partial u_0}\left(\frac{v_H(u_0)}{\mathbb{E}\left[e^{\varepsilon} v_L(u_0,\varepsilon)\right]}\right) > 0. \tag{26}$$

Equation (25) implies that the average elasticity of compensation with respect to idiosyncratic shock ε is increasing in promised utility u_0. 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, and $\frac{\partial \ln\left[e^{\varepsilon} c_L(u_0,\varepsilon)\right]}{\partial \varepsilon}$ is the elasticity of compensation with respect to idiosyncratic shock ε. Firms that promised a higher fraction of cash flow 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 would predict firm-level wage pass-throughs. In section 6.3, we show that, consistent with the above implication of our model, wage payments in firms with higher labor leverage are

[13] In our model the claim on aggregate dividends also has a higher price-to-dividend ratio in booms relative to recessions. In Appendix A2, we show that under Assumptions 1-5, $\exists\, \hat{\gamma} \in [1, 1+\xi)$ such that $\gamma > \hat{\gamma}$ implies $\frac{v_H(u_0^*)}{E\left[e^{\varepsilon} v_L(u_0^*,\varepsilon)\right]} > 1$.

more sensitive to firm-level idiosyncratic shocks.

Equation (26) summarizes the implications of our model on 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.

5 Quantitative Analysis

5.1 Numerical Algorithm

Policy functions and state prices depend on the infinite-dimensional state variable ϕ. The distribution ϕ directly shows up in the market clearing condition and indirectly as an argument in the stochastic discount factor in the description of the 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 state prices by projecting the stochastic discount factor on the space spanned by $\{g_t, x_t\}$ and use $x_{t+1} = \Gamma_x(g_{t+1}|x_t, g_t)$ 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 $\Gamma_x(x'|x, g)$ 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 A3, we describe a procedure that uses a grid on $\underline{\varepsilon}(u, g'|\phi, g, B)$, 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 update the law of motion Γ_x using simulated data. We repeat this procedure until the the function Γ_x converges. Appendix A3 describes the detailed steps

[14] The market clearing condition equation (16) implies that x is a $c(u|g, \phi, B)$ policy function 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.

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 4 years for recessions. The parameters for aggregate shocks are listed in the top part of table 1.

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. We use the specification $A(\theta) = a\left[\ln\left(\frac{1}{1-\theta}\right) - \theta\right]$ for the cost of retention effort. We interpret a separation in the model as a transition to the state of long-term unemployment (12 months and beyond). The parameters $\{a, \chi, \lambda, b\}$ are pinned down by the transition rates from employment to long-term unemployment, the duration of long-term unemployment, the average earnings

[15] 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. We use a more flexible process than the one listed section 2.1 to better fit the autocorrelation of aggregate consumption growth.

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.[17] We assume that both the Gaussian distributions as well as the negative exponential distribution satisfy a normalization 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.

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).[18] 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

[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%.

[17]The form of the negative exponential distribution is described in equation (A2.1) in Appendix A2.

[18]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 continuously, 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.

are then averaged separately for “boom years” and “recession years.”[19] We report the parameter values and moments related to the earnings distribution in the bottom part of table 1.

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 to -3.2% and -9%, respectively, in the PSID.[20]

All parameters affect aggregate labor shares. 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

We summarize aggregate asset pricing moments in table 2. 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 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 3.40% per year in the baseline model. In the data, the average debt-to-equity ratio for publicly traded U.S. firms is about 50%.[21] Accounting for this financial leverage, our model implies a market equity premium of 5.1%, which is close to the historical average excess return of 6.06% on the U.S.

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

[20]In a previous version, we also reported results for an alternative calibration which 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.

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

aggregate stock market index. In contrast, the equity premium is 0.39% 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 3, 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 38%, 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 capital owners' share of aggregate consumption x_t. Prolonged recessions are associated with low levels of the capital 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 capital owners. In equilibrium, the amplified volatility of the capital 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, higher return volatility, and predictability of market returns by valuation ratios. The model delivers a 9.35% standard 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 the risk-free rate, most of the increase in the volatility of the market return is accounted for by the time-varying equity premium.

Time variation in the risk premium also generates the predictability of future excess returns by price-to-dividend ratios, an empirical fact 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 "Model-Baseline" column report 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 "Data" column. Overall, the model produces regression coefficients and R^2 that are fairly close to 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 very low R^2.

Model benchmarking In this section, we compare our results to nested cases that capture important benchmarks in the literature. The comparisons highlight features of our model that are responsible for the quantitative results. 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 "Only-Worker-Side Limited Commitment" 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 change significantly 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-market stochastic discount factor and exogenous wage rigidity which 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 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 0.43% and a small volatility of excess returns of 2.77%. These values are only slightly higher than those in our first-best case reported under the column labeled "No Frictions" in table 2. 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 contrast, 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 the aggregate dividend claim. For example, under the column labeled "Model-Baseline" in table 2, while the risk premium on the aggregate consumption claim is 3.40%, the premium on the claim to corporate dividends is 3.53%. 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 the column labeled "No Mixture" in table 4, 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, $\sigma_H = \sigma_L$, we find that the asset pricing implications are almost similar to the first-best case, consistent with Krueger and Lustig (2010) intuition outlined in section 4.1. In the case $\sigma_L > \sigma_H$, it is possible to make σ_L to be 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 a equity premium on the unlevered aggregate consumption claim of 3.28%. 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.

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 value 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 are promised to workers. Firms with high-u workers have a high operating leverage and a low valuation ratio. Proposition 4 states that such firms should have a higher expected return. To our 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.

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 return 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-

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

month T-bill rate, on firm-level labor shares and time fixed effects.

$$\text{Excess Return}_{f,t+1} = \alpha_r + \beta_r \times \text{LaborShare}_{f,t} + \lambda_{rt}. \tag{27}$$

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

$$\text{LaborShare}_{f,t} = \frac{XLR_{ft}}{OPID_{f,t} + XLR_{f,t} + \Delta INV_{f,t}}, \tag{28}$$

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 (27), 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.

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 (14) 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. (forthcoming) 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.

[23] The estimates are robust to including various control variables such as leverage and total assets in the regression (27). See Appendix A5.

In our model, average separation rates are countercyclical: 3% per year in recessions and 2% per year in booms. Furthermore, most separations occur in worker-firm pairs where the value of the match is low. Relative to a mean separation rate of 2.2%, the separation rate in large and more productive firms is much smaller, about 0.5% per year. Endogenous separations mean that tail risk in labor earnings is partly driven 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% is 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.

Exposures to idiosyncratic and aggregate shocks Propositions 3 and 4 have direct implications on how idiosyncratic and aggregate shocks are insured in the presence of agency frictions. Workers with adverse histories are more exposed to idiosyncratic shocks in recessions due to firm-side limited commitment. 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

$$\begin{aligned}\Delta \log \text{WageBill}_{f,t+1} = \alpha_w + \beta_{w0}\ \text{LaborShare}_{f,t} + \beta_{w1}\Delta \log \text{Sales}_{f,t} \\ + \gamma_w \Delta \log \text{Sales}_{f,t} \times\ \text{LaborShare}_{f,t} + \lambda_{wt},\end{aligned} \tag{29}$$

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 (28). 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 1.[24] Furthermore, the interaction term $\gamma_w > 0$ and is statistically significant. This confirms the conclusion of Proposition 4 that firms with higher labor leverage have a higher pass-through coefficient. In Appendix A5, we estimate a version of (29) 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 mainly driven by the negative part of sales growth.

The history dependence of earnings' exposure to idiosyncratic shocks generated by the optimal contract also contrasts our model with models with exogenously incomplete markets; see for example, Constantinides and Duffie (1996), Schmidt (2015), and Constantinides and Ghosh (2014). In order to ensure a tractable equilibrium with zero trade in financial assets, these papers assume that the earnings processes of all workers have the same exposure to idiosyncratic shocks.

Our model has two related implications for how the exposure to aggregate shocks varies in the cross section. First, since the optimal risk sharing requires that high-productivity workers and capital owners insure low-productivity workers against aggregate shocks, their consumption will be more procyclical. Second, the risk-sharing scheme can be implemented by a portfolio strategy where high-productivity workers invest a higher fraction of their wealth in the aggregate stock market.

In the model, the standard deviation of consumption growth for capital 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 wealthiest 50% of stock-holding households is 7.8% and that of the wealthiest 75% is about 12% per year.

Next, we test the implication on wealth exposures. To illustrate the positive empirical relationship between wealth and stock market participation we run the following regression using the data from the 2007 Survey of Consumer Finances (SCF):[25]

$$\text{StockWeights}_i = \alpha + \beta \times \text{FinWealth Percentile}_i, \tag{30}$$

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

[25] This relationship is also documented by several other papers including Mankiw and Zeldes (1991), Poterba (2000), Vissing-Jorgensen and Attanasio (2003).

where StockWeights is defined as

$$\frac{\text{Stock Wealth}}{\text{Stock Wealth + Bond Wealth}}$$

and FinWealth Percentile is the percentile of Stock Wealth + Bond Wealth.[26] We find an intercept of 0.06 (s.e. 0.004) and a positive slope of 0.99 (s.e 0.009), confirming an increase in the stock market exposure with wealth.

In the model, equilibrium consumption of any agent can be replicated by a claim to the aggregate stock market index, a one-period risk-free bond, and a financial security whose payoff depends only on firm- and individual-level shocks, but not on aggregate shocks. We define $\Delta_C(u_i|g,\phi)$ to be the value of the aggregate stock market index as a fraction of the total value of worker i's consumption replicating portfolio. The details for the calculation of $\Delta_C(u|g,\phi)$ are in Appendix A4. We estimate regression (30) using data simulated by the model where we use $\Delta_C(u|g,\phi)$ as a proxy for StockWeights and human capital (which summarizes past idiosyncratic shocks) as a proxy for wealth. We find an intercept of 0.27 and a positive slope of 0.78.

The positive relationship between wealth and aggregate risk exposure contrasts us with models of only-worker-side limited commitment, for example, Alvarez and Jermann (2001). In these settings, (rich) workers who experienced a history of positive shocks are more likely to be constrained, and (poor) workers who experienced a history of negative shocks are more likely to be unconstrained. If the only-worker-side limited commitment model generates an amplified equity premium, then the discount factor of the marginal agent necessarily needs to be more volatile than that of an average agent. Interpreting such an insurance arrangement from the perspective of the consumption-replicating portfolio, unconstrained poor agents must have a higher weight in stocks than the constrained high-productivity workers; an implication that is inconsistent with the empirical evidence discussed above.

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

[26]To measure stock holdings, we sum direct holdings of equities and indirect holdings through mutual funds and retirement accounts. To measure bond holdings, we sum direct and indirect holdings of government bonds through mutual funds (taxable and nontaxable), saving bonds, liquid assets, money market accounts, and components of retirement accounts that are invested in government bonds. We restrict the sample to households with nonzero FinWealth and labor income.

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 observations on how individual earnings and wealth vary in their exposure to idiosyncratic and aggregate shocks.

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

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 (2014): "Liquidity and Risk Management: Coordinating Investment and Compensation Policies," *Working Paper, Columbia University*.

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, New York University*.

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.

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 Consumptionrisk," *Working Paper, University of Chicago Booth School of Business*.

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*, 2011, 1–72.

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

Fagereng, A., L. Guiso, and L. Pistaferri (2016): "Portfolio Choices, Firm Shocks and Uninsurable Wage Risk," *NBER Working 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.

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 (2014): "The Origins of Stock Market Fluctuations," *NBER Working Paper*.

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

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*, XLIX, 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., E. Pastorino, and V. Midrigan (forthcoming): "Debt Constraints and Employment," *Journal of Political Economy*.

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

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. A. Smith (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 M. SAI (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.

LUSTIG, H. N. AND S. G. VAN NIEUWERBURGH (2005): "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective," *The Journal of Finance*, 60, 1167–1219.

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

MANKIW, N. G. AND S. P. ZELDES (1991): "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics*, 29, 97–112.

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

POTERBA, J. M. (2000): "Stock Market Wealth and Consumption," *Journal of Economic Perspectives*, 14, 99–118.

RAMPINI, A. A. AND S. VISWANATHAN (forthcoming): "Financial Intermediary Capital," *Review of Economic Studies*, rdy020.

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

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

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

SHIMER, R. (2008): "The Probability of Finding a Job," in *American Economic Review*, vol. 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., R. LUCAS, JR., AND E. 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.

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

Vissing-Jorgensen, A. and O. P. Attanasio (2003): "Stock-market Participation, Intertemporal Substitution, and Risk-aversion," *American Economic Review*, 93, 383–391.

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

Figure 1: Timing of the two period model

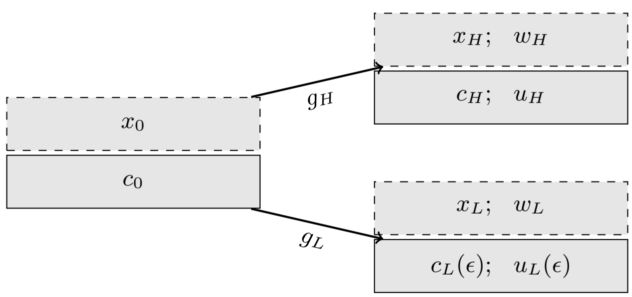

Table 1: PARAMETERS

Parameters	Values	Targeted moments	Values
Aggregate Risk			
g_H, g_L	0.35%, -0.15%	Mean, std of consumption growth	1.94%, 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	0.17	Mean separation rate	2.2%
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	0.01	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.

Table 2: AGGREGATE ASSET PRICING IMPLICATIONS

Moments	Model		Data
	Baseline	No Frictions	
Excess return on consumption			
mean	3.40%	0.39%	-
std.	8.87%	2.62%	-
Excess return on dividends			
mean	3.53%	0.43%	6.06%
std.	9.35%	2.77%	19.8%
Std of log SDF			
booms	17.52%	13.47%	38.00%
recessions	33.75%	20.07%	66.00%
Risk free rate			
mean	1.91%	4.73%	0.40%
std.	2.76%	0.39%	2.85%

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).

Table 3: AGGREGATE RETURN PREDICTABILITY

Horizon (quarters)	Model				Data	
	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).

Table 4: COMPARISON TO OTHER BENCHMARKS

Moments	Baseline	Only-Worker-Side Limited Commitment	Exogenous Wage Rigidity	No Mixture	
				$\sigma_H = \sigma_L$	$\sigma_H < \sigma_L$
Excess return on consumption					
mean	3.40%	1.24%	0.39%	1.07%	1.42%
std.	8.87%	3.60%	2.62%	3.15%	3.28%
Excess return on dividends					
mean	3.53%	1.17%	0.43%	1.10%	1.46%
std.	9.35%	3.36%	2.77%	3.34%	3.51%
Std of log SDF					
booms	17.52%	14.18%	13.47%	13.66%	17.92%
recessions	33.75%	22.07%	20.70%	23.67%	31.80%
Risk free rate					
mean	1.91%	3.58%	4.73%	4.28%	4.68%
std.	2.76%	0.72%	0.39%	0.74%	0.39%

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. For all cases, technology and preferences parameters are the same as the baseline. The column labeled "Only-Worker-Side Limited Commitment" relaxes constraint $v(u|\phi, B, g) \geq 0$. 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. The cash flow from this claim is modeled as $\tilde{x}(g)Y$ where $\tilde{x}(g)$ has a mean of 33% and a standard deviation of 2%, as in Favilukis and Lin (2016b). In the column labeled "No Mixture", we set the mixture probability of drawing from the negative exponential ρ to zero. The choices for values for $\{\sigma_H, \sigma_L\}$ in the subcolumns are explained in the text.

Table 5: FIRM-LEVEL RETURNS AND LABOR SHARES

Coefficients	Using LS	Using ELS
Labor share	1.38 (0.41)	1.25 (0.19)
Time fixed effects	Yes	Yes

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 (28), 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.

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

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.

Online Appendix for *Asset Pricing with Endogenously Uninsurable Tail Risk*

A1 Proof for Propositions 1 and 2

A1.1 Characterization of equilibrium

In this section, to prepare for the proofs for Propositions 1 and 2, we provide a set of necessary and sufficient conditions that characterize the equilibrium. We first state a lemma that establishes that the equality constraint (14) can be replaced by an inequality constraint so that the optimal contracting problem *P1* is a standard convex programming problem.

Lemma 1. *Suppose $A'(\theta)$, $A''(\theta)$, and $A'''(\theta) > 0$ for all $\theta \in (0,1)$. The policy functions for the optimal contracting problem P1 in the main text can be constructed from the solution to the a convex programming problem described below*

$$v\left(u|\phi,g,B\right) = \max_{c,\theta,\{u'(s')\}} \left\{ \begin{array}{c} 1-c-A(\theta)+ \\ \kappa\theta \int \Lambda\left(g'|\phi,g,B\right) e^{g'+\eta'+\varepsilon'} v\left(u'\left(s'\right)|\phi',g',B'\right)\Omega(ds'|g) \end{array} \right\} \tag{A1.1}$$

$$s.t: \quad u \leq \left[\left(1-\beta\right)c^{1-\frac{1}{\psi}} + \beta m^{1-\frac{1}{\psi}}\left(u|\,\phi,g,B\right)\right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{A1.2}$$

$$v\left(u'\left(s'\right)|\,\phi',g',B'\right) \geq 0,\ for\ \ all\ \ s', \tag{A1.3}$$

$$u'(s') \geq \lambda\overline{u}\left(\phi',g',B'\right),\ for\ \ all\ \ s', \tag{A1.4}$$

$$A'(\theta) \leq \kappa \int \Lambda\left(g'|\phi,g,B\right) e^{g'+\eta'+\varepsilon'} v\left(u'\left(s'\right)|\phi',g',B'\right)\Omega(ds'|g). \tag{A1.5}$$

where

$$m\left(u|\,\phi,g,B\right) = \left\{\kappa \int e^{(1-\gamma)\left(g'+\eta'+\varepsilon'\right)} \left[\theta\left[u'\left(s'\right)\right]^{1-\gamma} + (1-\theta)\left[\lambda\overline{u}\left(\phi',g',B'\right)\right]^{1-\gamma}\right]\Omega(ds'|g)\right\}^{\frac{1}{1-\gamma}},$$

Proof. We label the above-stated maximization problem as *P2*. The assumption that $A'(\theta)$ is strictly convex means that (A1.2)-(A1.5) describe a convex set with a nonempty interior and the objective function (A1.1) is concave. Thus, problem *P2* is a convex programming problem. We next show that optimal choices for *P2* are feasible for problem *P1*.

Optimal policies for *P2* satisfy a set of first-order necessary conditions. In particular, let $\iota \geq 0$ be the Lagrange multiplier of the constraint (A1.5), first-order conditions with respect to θ implies

$$\iota A''\left(\theta\right) = \frac{\beta}{1-\beta} c^{\frac{1}{\psi}} m^{\gamma-\frac{1}{\psi}} \frac{1}{1-\gamma} \int e^{(1-\gamma)\left(\eta'+\varepsilon'\right)} \left\{\left[u'\left(s'\right)\right]^{1-\gamma} - \left[\lambda\overline{u}\left(g',\phi',B'\right)\right]^{1-\gamma}\right\} \Omega\left(ds'|\,g'\right). \tag{A1.6}$$

The limited commitment constraint on worker side, equation (A1.4) implies that right-hand side of (A1.6) must be strictly positive. Therefore, $\iota > 0$ and (A1.5) must holds with equality at the optimum.

Let ι_u be the Lagrange multiplier of the promise keeping constraint (A1.2), the first-order condition with respect to c implies

$$\iota_u = \frac{1}{1-\beta}\left(\frac{c}{u}\right)^{\frac{1}{\psi}} > 0. \tag{A1.7}$$

Thus, inequality (A1.2) must also hold with equality at the optimum.

As a result, the optimal policy for *P2* satisfy all of the constraints for *P1* and as the constraint set for *P2* larger, the optimal policies to *P2* also attain the maximum for *P1*.

∎

Suppose the stochastic discount factor and the law of motion of the aggregate state variables jointly satisfy the following condition:

Assumption A.1. *For some $\epsilon > 0$, and for all (ϕ, g, B),*

$$\sum \pi\left(g'|\,g\right)\Lambda\left(g'|\phi, g, B\right)e^{g'} < 1-\epsilon. \tag{A1.8}$$

Given Assumption A.1, standard arguments from Stokey et. al (1989) imply that there is a unique v in the space of bounded continuous functions that satisfies (A1.1). In addition, v is continuous, strictly decreasing, strictly concave and differentiable in the interior. We denote the policy functions for *P2* by

$$\left\{c\left(u|\,\phi, g, B\right), \theta\left(u|\,\phi, g, B\right), \iota(u|\phi, g, B), \{\bar{\varepsilon}(u, g'|\phi, g), \underline{\varepsilon}(u, g'|\phi, g)\}_{g'}, \{u'\left(u, z'|\,\phi, g, B\right)\}_{z'}\right\}. \tag{A1.9}$$

The first-order necessary conditions for *P2* imply that the above policy functions must satisfy

1. $\forall \eta' + \varepsilon' \in [\underline{\varepsilon}(u, g'|\phi, g, B),\ \bar{\varepsilon}(u, g'|\phi, g, B)]$, $u'\left(u, s'|\,\phi, g, B\right)$ satisfy

$$\Lambda\left(g'|\phi, g, B\right) = \frac{\beta e^{-\gamma\left(g'+\eta'+\varepsilon'\right)}}{1+\frac{\iota(u|\phi,g,B)}{\theta(u|\phi,g,B)}}\left[\frac{c\left(u'\left(u, s'|\,\phi, g, B\right)|\,\phi', g', B'\right)}{c\left(u|\,\phi, g, B\right)}\right]^{-\frac{1}{\psi}}\left[\frac{u'\left(u, s'|\,\phi, g, B\right)}{m\left(u|\,\phi, g, B\right)}\right]^{\frac{1}{\psi}-\gamma}. \tag{A1.10}$$

2. $\forall \eta' + \varepsilon' \geq \bar{\varepsilon}(u, g'|\phi, g, B)$,

$$u'\left(u, s'|\,\phi, g, B\right) = \lambda\bar{u}\left(g', \phi', B'\right); \tag{A1.11}$$

   and $\forall \eta' + \varepsilon' \leq \underline{\varepsilon}(u, g'|\phi, g, B)$,

$$u'\left(u, s'|\,\phi, g, B\right) = u^*\left(g', \phi', B'\right), \tag{A1.12}$$

   where $u^*\left(g', \phi', B'\right)$ satisfies that $v\left(u^*\left(g, \phi, B\right)|\,\phi, g, B\right) = 0$ for all (ϕ, g, B).

3. The Lagrange multiplier $\iota(u|\phi, g, B)$ satisfies

$$\begin{aligned}\iota(u|\phi, g, B) &= \frac{1}{A''\left(\theta\left(u|\,\phi, g, B\right)\right)}\frac{\beta}{1-\beta}c\left(u|\,\phi, g, B\right)^{\frac{1}{\psi}}m\left(u|\,\phi, g, B\right)^{\gamma-\frac{1}{\psi}} \times \frac{1}{1-\gamma} \\ &\times \left\{\int e^{(1-\gamma)\left(\eta'+\varepsilon'\right)}\left\{\left[u'\left(u, s'|\,\phi, g, B\right)\right]^{1-\gamma} - \left[\lambda\bar{u}\left(g', \phi', B'\right)\right]^{1-\gamma}\right\}\Omega\left(ds'|\,g'\right)\right\}.\end{aligned} \tag{A1.13}$$

   The policy functions must satisfy the equality constraints of the problem *P1*

$$A'(\theta\left(u|\,\phi, g, B\right)) = \kappa\int\Lambda\left(g'|\phi, g, B\right)e^{g'+\eta'+\varepsilon'}v\left(u'\left(s'\right)|\phi', g', B'\right)\Omega(ds'|g), \tag{A1.14}$$

$$u = \left[(1-\beta)\, c^{1-\frac{1}{\psi}} + \beta m^{1-\frac{1}{\psi}} \left(u \middle| \phi, g, B \right) \right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{A1.15}$$

where

$$m\left(u \middle| \phi, g, B \right) = \left\{ \kappa \int e^{(1-\gamma)\left(g'+\eta'+\varepsilon'\right)} \begin{bmatrix} \theta\left(u \middle| \phi, g, B \right) \left[u'\left(u, s' \middle| \phi, g, B \right) \right]^{1-\gamma} \\ +(1-\theta\left(u \middle| \phi, g, B \right)) \left[\lambda \overline{u}\left(\phi', g', B' \right) \right]^{1-\gamma} \end{bmatrix} \Omega(ds'|g) \right\}^{\frac{1}{1-\gamma}}.$$

The following lemma states that conditions (A1.10) - (A1.15) are both necessary and sufficient for optimality.

Lemma 2. *Suppose there exist an SDF $\Lambda\left(g'|\phi, g, B\right)$, a worker's outside option, $\bar{u}\left(\phi, g, B\right)$, and a law motion for aggregate state variables that satisfy Assumption A.1. Suppose that given $\Lambda\left(g'|\phi, g, B\right)$, $\bar{u}\left(\phi, g, B\right)$, and the law of motion for state variables, policy functions for problem P2, as denoted in (A1.9), satisfy the optimality conditions (A1.10)-(A1.13) and the equality constraints (A1.14)-(A1.15). In addition, $\frac{c(u|\phi,g,B)}{u}$ is nondecreasing in u for all (g, ϕ, B). Let $v\left(u|\phi, g, B\right)$ be the unique fixed point of the operator T:*

$$Tv\left(u|\phi, g, B\right) = \begin{array}{c} 1 - c\left(u \middle| \phi, g, B \right) - A(\theta\left(u \middle| \phi, g, B \right)) + \\ \kappa\theta\left(u \middle| \phi, g, B \right) \int \Lambda\left(g'|\phi, g, B\right) e^{g'+\eta'+\varepsilon'} v\left(u'\left(u, s' \middle| \phi, g, B \right) |\phi', g', B'\right) \Omega(ds'|g). \end{array} \tag{A1.16}$$

Then, the policy functions together with the value function $v(u|g, \phi, B)$ solve the problem P2.

Proof. Suppose there exists a set of policy functions that satisfy conditions (A1.10)-(A1.15). Given condition (A.1), the operator defined in (A1.16) is a contraction, and we can construct the value function $v\left(u \middle| \phi, g, B \right)$ from the policy functions as the unique fixed point of (A1.16). The first-order conditions (A1.10)-(A1.12) imply that the value function constructed above must satisfy

$$\frac{\partial}{\partial u} v\left(u \middle| \phi, g, B \right) = -\frac{1}{1-\beta} \left(\frac{c\left(u \middle| \phi, g, B \right)}{u} \right)^{\frac{1}{\psi}}. \tag{A1.17}$$

Because $\frac{c(u|\phi,g,B)}{u}$ is nondecreasing in u, $\frac{\partial}{\partial u} v\left(u \middle| \phi, g, B \right)$ must be nonincreasing, that is, $v\left(u \middle| \phi, g, B \right)$ is a concave function of u. As a result, given $v\left(u \middle| \phi, g, B \right)$, the first-order conditions, (A1.10)-(A1.15) can be shown to be equivalent to the set of first-order conditions for the programming problem *P2*, which is necessary and sufficient for optimality. Therefore, the above constructed value functions and policy functions must solve the optimal contracting problem *P2*, as needed. ∎

Given the above discussion, it is straightforward to provide a characterization for the equilibrium price and quantities using optimality conditions. We summarize these conditions in the following lemma. The proof is omitted as it follows directly from Lemma 1 and Lemma 2.

Lemma 3. *The equilibrium prices and quantities can be summarized as:*

i) *a set of policy functions,*

$$x\left(g, \phi, B\right), c\left(u \middle| \phi, g, B \right), \theta\left(u \middle| \phi, g, B \right), \iota(u|\phi, g, B), \left\{ \bar{\varepsilon}(u, g'|\phi, g), \underline{\varepsilon}(u, g'|\phi, g) \right\}_{g'}, \left\{ u'\left(u, s' \middle| \phi, g, B \right) \right\}_{s'},$$

ii) *worker' outside option $\bar{u}\left(\phi, g, B\right)$ and initial utility at employment $u^*\left(\phi, g, B\right)$,*

iii) *a law of motion of ϕ and B,*

iv) *a SDF and a firm value function $v\left(u \middle| \phi, g, B \right)$, such that*

1. *the SDF is consistent with capital owner's consumption, that is,* $\Lambda(g'|\phi, g, B)$ *and* $x(g, \phi, B)$ *satisfy equation (19), where the capital owner's utility,* $w(g, \phi, B)$ *is constructed from* $x(g, \phi, B)$ *using equation (15),*

2. *the value function and policy functions satisfy the optimality conditions (A1.10)-(A1.15),*

3. *the outside option* $\bar{u}(\phi, g, B)$ *satisfies (9),* $u^*(\phi, g, B)$ *satisfies* $v(u^*(g, \phi, B)|\phi, g, B) = 0$ *for all* (ϕ, g, B)*, and*

4. *the law of motion of the aggregate state variables satisfy (17) and (18).*

We now prove Proposition 1.

A1.2 Proof of Proposition 1

Given Assumption A.1 and Lemma 1, standard arguments from Stokey et. al (1989) imply that the value function v for the optimal contracting problem (10) is continuous, strictly decreasing, strictly concave and differentiable in the interior. Because the value function is strictly decreasing, the limited commitment constraint (12) can be written as $u'(s') \leq u^*(\phi', g', B')$ for all s', where $u^*(\phi, g, B)$ is defined by equation (20). Therefore, the first order condition with respect to continuation utility and the envenlop condition for the programming problem (A1.1) together imply that one of the following three cases have to true:

1. In the interior, equation (23) holds.

2. The worker-side limited commitment constraint binds, $u'(u, s'|\phi, g, B) = \lambda\bar{u}(\phi', g', B')$, and,

$$\begin{aligned} & \left[\frac{x(\phi', g', B')}{x(\phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{w(\phi', g', B')}{n(\phi, g, B)}\right]^{\frac{1}{\psi}-\gamma} \left(1 + \frac{\iota(u|\phi, g, B)}{\theta(u|\phi, g, B)}\right) \\ \geq \quad & e^{-\gamma(\eta'+\varepsilon')} \left[\frac{c(u'(u, s'|\phi, g, B), \phi', B')}{c(u|\phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{u'(u, s'|\phi, g, B)}{m(u|\phi, g, B)}\right]^{\frac{1}{\psi}-\gamma}, \end{aligned} \tag{A1.18}$$

3. The firm-side limited commitment costraint binds, $u'(s') = u^*(\phi', g', B')$,

$$\begin{aligned} & \left[\frac{x(\phi', g', B')}{x(\phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{w(\phi', g', B')}{n(\phi, g, B)}\right]^{\frac{1}{\psi}-\gamma} \left(1 + \frac{\iota(u|\phi, g, B)}{\theta(u|\phi, g, B)}\right) \\ \leq \quad & e^{-\gamma(\eta'+\varepsilon')} \left[\frac{c(u'(u, s'|\phi, g, B), \phi', B')}{c(u|\phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{u'(u, s'|\phi, g, B)}{m(u|\phi, g, B)}\right]^{\frac{1}{\psi}-\gamma}. \end{aligned} \tag{A1.19}$$

Define $\mathcal{E} = \{\eta' + \varepsilon' : \textit{ equation } (23) \textit{ holds}\}$. Also, let

$$\underline{\varepsilon}(u, g'|\phi, g, B) = \inf \mathcal{E}, \quad \bar{\varepsilon}(u, g'|\phi, g, B) = \sup \mathcal{E}. \tag{A1.20}$$

Let $l_u(u|\phi, g, B)$ be the Lagrange multiplier for the promise-keeping constraint of the programming problem (A1.1), then

$$\frac{\partial}{\partial u} v(u|\phi, g, B) = l_u(u|\phi, g, B) = \frac{1}{1-\beta} \left(\frac{c(u|\phi, g, B)}{u}\right)^{\frac{1}{\psi}}, \tag{A1.21}$$

where the first equality is the envelope theorem, and the second equality is the first order condition, (A1.7). Because v is concave, the above condition implies that $c(u|\phi, g, B)$ must be strictly increasing in u. Thereore, the optimality condition (23) implies that on $\mathcal{E}$, $u'(u, g', \eta', \varepsilon' | \phi, g, B)$ must be strictly decreasing in $\eta' + \varepsilon'$. Clearly, the strict monotonicity of $u'(u, g', \eta', \varepsilon' | \phi, g, B)$ implies that $u'(u, g', \eta', \varepsilon' | \phi, g, B) = \lambda \bar{u}(\phi', g', B')$ if $\eta' + \varepsilon' = \underline{\varepsilon}(u, g' | \phi, g, B)$ and $u'(u, g', \eta', \varepsilon' | \phi, g, B) = u^*(\phi', g', B')$ if $\eta' + \varepsilon' = \overline{\varepsilon}(u, g' | \phi, g, B)$.

First, $\forall \eta' + \varepsilon' < \underline{\varepsilon}(u, g' | \phi, g, B)$, we must have $u'(u, g', \eta', \varepsilon' | \phi, g, B) = \lambda \bar{u}(\phi', g', B')$. Otherwise, none of the equations, (23), (A1.18), or (A1.19) can hold. Similarly, $\forall \eta' + \varepsilon' > \overline{\varepsilon}(u, g' | \phi, g, B)$, we must have $u'(u, g', \eta', \varepsilon' | \phi, g, B) = u^*(\phi', g', B')$.

Second, to complete the proof of Part 1 and 2 of Proposition 1, we need to show that $\forall \eta' + \varepsilon' \in (\underline{\varepsilon}(u, g' | \phi, g, B),\ \overline{\varepsilon}(u, g' | \phi, g, B))$, condition (23) must hold. It is enough to show $u'(u, g', \eta', \varepsilon' | \phi, g, B) \in (\lambda \bar{u}(\phi', g', B'),\ u^*(\phi', g', B'))$. This can be proved by contradiction. Suppose $\eta' + \varepsilon' \in (\underline{\varepsilon}(u, g' | \phi, g, B),\ \overline{\varepsilon}(u, g' | \phi, g, B))$ and $u'(u, g', \eta', \varepsilon' | \phi, g, B) = \lambda \bar{u}(\phi', g', B')$, then the fact that equation (23) holds at $\overline{\varepsilon}(u, g' | \phi, g, B)$ implies that (note that $\eta' + \varepsilon' < \overline{\varepsilon}(u, g' | \phi, g, B)$)

$$
\begin{aligned}
&\left[\frac{x(\phi', g', B')}{x(\phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{w(\phi', g', B')}{n(\phi, g, B)}\right]^{\frac{1}{\psi} - \gamma} \left(1 + \frac{\iota(u | \phi, g, B)}{\theta(u | \phi, g, B)}\right) \\
< \quad & e^{-\gamma(\eta' + \varepsilon')} \left[\frac{c(\lambda \bar{u}(\phi', g', B'), \phi', B')}{c(u | \phi, g, B)}\right]^{-\frac{1}{\psi}} \left[\frac{\lambda \bar{u}(\phi', g', B')}{m(u | \phi, g, B)}\right]^{\frac{1}{\psi} - \gamma},
\end{aligned}
$$

which is a contradiction to condition (A1.18). Similarly, one can show that $u'(u, g', \eta', \varepsilon' | \phi, g, B) = u^*(\phi', g', B')$ cannot be true either.

To prove Part 3 of Proposition 1, note that because the value function is strictly concave in u, the Lagrange multiplier $\iota_u(u | \phi, g, B)$ must be strictly increasing in u. The first-order condition with respect to $u'(u, g', \eta', \varepsilon' | \phi, g, B)$ in the programming problem (A1.1) then implies that $u'(u, g', \eta', \varepsilon' | \phi, g, B)$ must be strictly increasing in u as well. Given constraint (14), the monotonicity of $\theta(u | \phi, g, B)$ with respect to u then follows directly from the $u'(u, g', \eta', \varepsilon' | \phi, g, B)$ is increasing with respect to u and the fact that $v(u' | \phi', g', B')$ is strictly decreasing in u'.

A1.3 Proof of Proposition 2

Proposition 2 follows directly from the lemma below, which provides the details for the construction of the equilibrium in the stochastic economy from a given equilibrium of the deterministic economy with a modified discount rate.

Lemma 4. *Suppose g_t is i.i.d. over time and $f(\cdot | g)$ does not depend on g. Suppose there exists an equilibrium in the equivalent deterministic economy with modified discount rate. An equilibrium of the stochastic economy can be constructed as follows.*

i) *The SDF is given by equation (24) in Proposition 2.*

ii) *Workers' outside option and utility upon employment are given by:*

$$
\bar{u}(\phi, g, B) = \widehat{\bar{u}}(\phi, B), \quad u^*(\phi, g, B) = \widehat{u}^*(\phi, B),
$$

respectively, where $\widehat{\bar{u}}(\phi, B)$ and $\widehat{u}^*(\phi, B)$ are the corresponding equilibrium quantities in the equivalent deterministic economy with a modified discount rate.

iii) *The consumption share of capital owners is*

$$
x(\phi, g, B) = \hat{x}(\phi, B),
$$

where $\hat{x}(\phi, B)$ is the capital owner's consumption share in the equivalent deterministic economy with a modified discount rate.

iv) *The value function and policy functions of the optimal contracting problem are given by*

$$v(u|\phi, g, B) = \hat{v}(u|\phi, B), \; c(u|\phi, g, B) = \hat{c}(u|\phi, B),$$
$$\theta(u|\phi, g, B) = \hat{\theta}(u|\phi, B), \; u'(u, g', \varepsilon'|\phi, g, B) = \hat{u}'(u, \varepsilon'|\phi, B).$$

v) *The law of motion for aggregate state variables (ϕ, B) is the same as that in the equivalent deterministic economy with a modified discount rate.*

Proof. To prove that the proposed allocations and prices constitutes an equilibrium, we use Lemma 3 to verify the equilibrium conditions. First, we show that the proposed stochastic discount factor is consistent with capital owners' consumption and utility process. Given capital owner's consumption and utlity in the stochastic economy, using equation (19),

$$\Lambda(g'|\phi, g, B) = \beta e^{-\gamma g'} \left[\frac{\hat{x}(\phi', B')}{\hat{x}(\phi, B)}\right]^{-\frac{1}{\psi}} \left[\frac{w(\phi', B')}{\left(\mathbb{E}\left[e^{(1-\gamma)g'}\hat{w}^{1-\gamma}(\phi', B')\right]\right)^{\frac{1}{1-\gamma}}}\right]^{\frac{1}{\psi}-\gamma}$$

The utility $w(\phi', B')$ is deterministic and the above can be written as

$$\Lambda(g'|\phi, g, B) = \beta e^{-\gamma g'} \left[\frac{\hat{x}(\phi', B')}{\hat{x}(\phi, B)}\right]^{-\frac{1}{\psi}} \left(E\left[e^{(1-\gamma)g'}\right]\right)^{\frac{\gamma-\frac{1}{\psi}}{1-\gamma}} = \hat{\beta}\left[\frac{\hat{x}(\phi', B')}{\hat{x}(\phi, B)}\right]^{-\frac{1}{\psi}} \frac{e^{-\gamma g'}}{\mathbb{E}\left[e^{(1-\gamma)g'}\right]}. \quad \text{(A1.22)}$$

Given the consumption policy in the deterministic economy, the SDF in the deterministic economy reduces to a risk-free discount rate $R(\phi, B)$ with

$$\frac{1}{R(\phi, B)} = \hat{\beta}\left[\frac{x(\phi', B')}{x(\phi, B)}\right]^{-\frac{1}{\psi}}. \quad \text{(A1.23)}$$

Combing equations (A1.22) and (A1.23), it is clear that the SDF defined in (24) is consistent with capital owners' consumption in the stochastic economy.

Next, we show that the proposed value function and policy functions also solve the optimal contracting problem in the stochastic economy. It is enough to show that the value function of the deterministic economy is also a fixed point of the Bellman operator implied by the optimal contracting problem *P1*. Given that the two economies have the same workers' outside options $\bar{u}(\phi, g, B)$ and that $\hat{\beta} = \beta\left(\mathbb{E}\left[e^{(1-\gamma)g'}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}$, it is easy to see that constraints (11), (12), and (13) are identical in both economies. Given $v(u|\phi, B)$ the term

$$\int \Lambda(g'|\phi, g, B)\, e^{g'+\eta'+\varepsilon'} v(u'(s')|\phi', g', B')\, \Omega(ds'|g)$$

can be written as

$$\begin{aligned}
& \frac{1}{R(\phi, B)} \sum \pi(g') \int \frac{e^{-\gamma g'}}{\mathbb{E}\left[e^{(1-\gamma)g'}\right]} e^{g'+\varepsilon'+\eta'} v(u'(\varepsilon', \eta') | \phi', B') f(\varepsilon' + \eta') d\varepsilon' d\eta' \\
= & \frac{1}{R(\phi, B)} \int \frac{\mathbb{E}\left[e^{(1-\gamma)g'}\right]}{\mathbb{E}\left[e^{(1-\gamma)g'}\right]} e^{\varepsilon'+\eta'} v(u'(\varepsilon', \eta') | \phi', B') f(\varepsilon' + \eta') d\varepsilon' d\eta' \\
= & \frac{1}{R(\phi, B)} \int e^{\varepsilon'+\eta'} v(u'(\varepsilon', \eta') | \phi', B') f(\varepsilon' + \eta') d\varepsilon' d\eta',
\end{aligned}$$

which is identical to that in the deterministic economy. Therefore, $v(u|\phi, g, B) = \hat{v}(u|\phi, B)$ is also the value function of the optimal contracting problem in the stochastic economy.

Finally, conditions 3 and 4 in Lemma 3 also hold, because these requirements are identical in the deterministic economy and the stochastic economy. This completes the proof.

∎

A2 Proofs of Propositions 3 and 4

A2.1 Equilibrium in the simple economy

In this section, we start with deriving explicit expressions for several equilibrium objects to prepare for the proofs of Propositions 3 and 4. We first introduce some notation.

Notation In the simple model in Section 4, we assume that the worker-specific shock follows a negative exponential distribution. The density of a negative exponential distribution with parameter ξ takes the following form:

$$f(\varepsilon | g_L) = \begin{matrix} 0 & \varepsilon > \varepsilon_{MAX} \\ \xi e^{\xi(\varepsilon - \varepsilon_{MAX})} & \varepsilon \leq \varepsilon_{MAX}. \end{matrix} \tag{A2.1}$$

For later reference, we note that the moments of $f(\varepsilon | g_L)$ can be easily computed as

$$\int_{-\infty}^{\varepsilon} e^{\theta t} f(t | g_L) dt = \frac{\xi}{\xi + \theta} e^{-\xi \varepsilon_{MAX} + (\theta + \xi)\varepsilon} \quad for \quad \xi + \theta > 0. \tag{A2.2}$$

Equation (A2.2) shows that the assumption $\mathbb{E}[e^{\varepsilon}] = 1$ amounts to a parameter restriction that $\varepsilon_{MAX} = \ln \frac{1+\xi}{\xi}$.

In the simple economy illustrated in figure 1, we let $x_H \equiv x(g_H)$ and $x_L \equiv x(g_L)$ denote capital owners' consumption share and $w_H \equiv w(g_H)$ and $w_L \equiv w(g_L)$ denote their normalized utility at node H and L, respectively.

In solving the optimal contracting problem, it is more convenient to represent policy functions and value functions as functions of the period-0 promised utility u_0. For an arbitrary u_0, we use $u_H(u_0) \equiv u'(u_0, g_H)$, and $u_L(u_0, \varepsilon') \equiv u'(u_0, g_L, \varepsilon')$ to denote the normalized promised utility for a worker with initial promised utility u_0 at nodes H and L, respectively. We use $c_0(u_0)$, $c_H(u_0) \equiv c(u_H(u_0)| g_H)$, and $c_L(u_0, \varepsilon') = c(u_L(u_0, \varepsilon')|, g_L)$ for workers' consumption policy at nodes 0, H, and L, respectively. Similarly, $v_H(u_0) \equiv v(u_H(u_0)| g_H)$, $v_L(u_0, \varepsilon') \equiv v(u_L(u_0, \varepsilon')| g_L)$, $\theta_H(u_0) \equiv \theta(u_H(u_0)| g_H)$ and $\theta_L(u_0, \varepsilon') \equiv \theta(u_L(u_0, \varepsilon')| g_L)$ are value functions and policy functions at note H and L, respectively. We also denote $\underline{\varepsilon}_L(u_0) \equiv \underline{\varepsilon}(u_0, g_L)$ as the lowest level of realization of the ε' shock such that the limited commitment constraint does not bind at node L.

In addition, let u_H^{FB} and u_L^{FB} denote the utility-to-consumption ratio of an agent who consumes the aggregate consumption in state g_H and g_L, respectively. That is, they are the normalized utility associated with full risk sharing. The first best levels, u_H^{FB} and u_L^{FB} are determined by

$$u_H^{FB} = \left(e^{g_H} u_H^{FB}\right)^{\beta} \quad u_L^{FB} = \left(e^{g_L} u_L^{FB}\right)^{\beta}.$$

Also, we use u_L^{CD} to denote the normalized utility of an agent in an economy without risk sharing. That is, it is utility-consumption ratio of an agent who consumes y_t every period:

$$u_L^{CD} = \left(\int \left[e^{\{\varepsilon' + g_L\}} u_L^{CD}\right]^{1-\gamma} f(\varepsilon'|g_L) d\varepsilon\right)^{\frac{\beta}{1-\gamma}}. \tag{A2.3}$$

It is straightforward to show that as $\gamma \to 1 + \xi$, $u_L^{CD} \to 0$. We solve the general equilibrium in the simple economy by backward induction. We first solve the value functions and policy functions at nodes H and L in period 1. In the second step, we analyze the optimal contracting problem in period 0 for an arbitrary promised utility u_0. Finally, we impose market clearing to solve for the equilibrium stochastic discount factor.

Value functions at nodes H and L The following lemma characterizes the value functions at nodes H and L in period 1.

Lemma 5. *(Value function in period 1)*

The firm's value function at nodes H and L are give by

$$v\left(u|\,g_H\right) = 1 - c\left(u|\,g_H\right) + \frac{\beta}{1-\beta} x_H - a \ln\left[1 + \frac{\beta x_H}{a\left(1-\beta\right)}\right], \text{ and} \tag{A2.4}$$

$$v_L\left(u|\,g_L\right) = 1 - c\left(u|\,g_L\right) + \frac{\beta}{1-\beta} x_L - a \ln\left[1 + \frac{\beta x_L}{a\left(1-\beta\right)}\right], \tag{A2.5}$$

respectively, where the consumption policies are given by

$$c\left(u|\,g_H\right) = \left(\alpha e^{g_H} u_H^{FB}\right)^{-\frac{\beta}{1-\beta}} u^{\frac{1}{1-\beta}}, \quad c\left(u|\,g_L\right) = \left(\alpha \Upsilon e^{g_L} u^{CD}\right)^{-\frac{\beta}{1-\beta}} u^{\frac{1}{1-\beta}},$$

and the effort choices are

$$\theta_H = 1 - \frac{a}{a + \frac{\beta}{1-\beta} x_H}, \quad \theta_L = 1 - \frac{a}{a + \frac{\beta}{1-\beta} x_L}. \tag{A2.6}$$

Proof. Here, we only provide details for the deviation of the value function at node H. The value function at node L can be computed in the same way. At node H, the optimal contracting problem is written as

$$v\left(u|\,g_H\right) = \max_{c,\theta,u'} \left\{1 - c - A\left(\theta\right) + \theta \frac{1}{R_H} e^{g_H} v_{2,H}\left(u'\right)\right\} \tag{A2.7}$$

$$\text{subject to: } u = c^{1-\beta}\left(e^{g_H} u'\right)^{\beta}$$

$$u' = \alpha u_H^{FB}$$

$$v_{2,H}\left(u'\right) \geq 0$$

$$A'\left(\theta\right) = \frac{1}{R_H} e^{g_H} v_{2,H}\left(u'\right)$$

where we use $v_{2,H}\left(u'\right)$ for the value function in period 2. Because there is no aggregate uncertainty in period 2, we replace the stochastic discount factor by a risk-free discount rate, $\frac{1}{R_H}$. The absence of idiosyncratic shocks and the fact that workers consume α fraction of their output imply that workers' utility is α times the utility of a representative consumer, that is $u' = \alpha u_H^{FB}$. Note also, because the firm always receive $1-\alpha$ fraction of y_t after period 2, the limited commitment constraint $v_{2,H}(u') \geq 0$ does not bind.

To derive a close-form solution for $v_H(u)$, we first note that $v_{2,H}(u') = \frac{1-\alpha}{1-\beta}$. From period 2 and on, capital owner's consumption and firms' cash flow are both proportional to aggregate output. Under the assumption of unit elasticity, the price-to-dividend ratio of the firm's cash flow is $\frac{1}{1-\beta}$. Because the firm receive $1-\alpha$ fraction of y_t, the ratio of firm value normalized by y_t is $\frac{1-\alpha}{1-\beta}$.

Second, because capital owner's consumption share is x_H in period 1 and $1-\alpha$ in period 2, the discount factor is $\frac{1}{R_H} = \beta\left[\frac{1-\alpha}{x_H}e^{g_H}\right]^{-1}$. Therefore, the value function can be written as

$$v_H(u) = 1 - c - A(\theta) + \theta\frac{\beta}{1-\beta}x_H. \tag{A2.8}$$

The consumption policy can be backed out from the promise-keeping constraint $u = c^{1-\beta}\left(e^{g_H}u'\right)^{\beta}$. In addition, given then functional form of $A(\theta)$, the optimal effort θ can be solved from the incentive constraint, $A'(\theta) = \frac{\beta}{1-\beta}x_H$, which gives (A2.6). Replacing θ in (A2.8) with the optimal policy, we obtain the representation of the value function in (A2.4).

The optimal contracting problem at node L has a similar structure:

$$\begin{aligned}
v(u|\,g_L) = \max_{c,\theta,u'(\varepsilon')} &\left\{1 - c - A(\theta) + \theta\frac{1}{R_L}\int_{-\infty}^{\infty} e^{g_L+\varepsilon'}v_{2,L}\left(u'(\varepsilon')\right)f(\varepsilon'|\,g_L)\,d\varepsilon'\right\}\\
\textit{subject to}:\ u &= c^{1-\beta}\left\{\int_{-\infty}^{\infty}\left[e^{g_L+\varepsilon'}u'(\varepsilon')\right]^{1-\gamma}f(\varepsilon'|\,g_L)\,d\varepsilon'\right\}^{\frac{\beta}{1-\gamma}}\\
u'(\varepsilon') &= \alpha u_L^{CD}\\
v_{2,L}\left(u'(\varepsilon')\right) &\geq 0\\
A'(\theta) &= \frac{1}{R_L}e^{g_L}\int_{-\infty}^{\infty}v_{2,L}\left(u'(\varepsilon')\right)f(\varepsilon'|\,g_L)\,d\varepsilon'.
\end{aligned}$$

The above problem can be greatly simplified by noting that $v_{2,L}(u'(\varepsilon')) = \frac{1-\alpha}{1-\beta}$ and $u'(\varepsilon') = \alpha u_L^{CD}$ do not depend on ε'. Also, we define

$$\Upsilon = \left\{\int_{-\infty}^{\infty}e^{(1-\gamma)\varepsilon'}f(\varepsilon'|\,g_L)\,d\varepsilon'\right\}^{\frac{1}{1-\gamma}}, \tag{A2.9}$$

so that $\left\{\int_{-\infty}^{\infty}\left[e^{\varepsilon'}u'(\varepsilon')\right]^{1-\gamma}f(\varepsilon'|\,g_L)\,d\varepsilon'\right\}^{\frac{1}{1-\gamma}} = \alpha u_L^{CD}\Upsilon$. The rest of the proof can be completed by following the same steps in the solution of (A2.7). ∎

At node L, limited commitment on firm side requires that $v_L(u) \geq 0$. Therefore, by equation (A2.5), the maximum amount of consumption that the firm can promise to deliver to a worker at node L is $1 - A(\theta_L) + \theta_L\frac{\beta}{1-\beta}x_L$, which we will denote as c_L^{MAX}. Recall that for a worker with initial promised utility u_0, $\underline{\varepsilon}_L(u_0)$ is the lowest level of realization of the ε' shock such that the limited commmitment constraint does not bind at node L. We must have, for all u_0,

$$c_L\left(u_0, \underline{\varepsilon}_L(u_0)\right) = 1 + \frac{\beta}{1-\beta}x_L - a\ln\left[1 + \frac{\beta x_L}{a(1-\beta)}\right]. \tag{A2.10}$$

We now turn to the optimal contracting problem as node 0.

Optimal contracting at node 0 We first prove the following lemma that uses the optimal risk sharing condition (23) to relate the marginal rate of substitution of a marginal worker whose limited commitment constraint is just about to bind to that of the capital owners.

Lemma 6. *(FOC for the marginal agent)*

Given the consumption share of the capital owners, x_H and x_L, for all u_0, the normalized consumption of the marginal worker with $\varepsilon_1 = \underline{\varepsilon}_L(u_0)$ must satisfy:

$$\frac{c_H(u_0)}{e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L(u_0, \underline{\varepsilon}_L(u_0))} \left[\frac{u_L^{FB} k(\theta_H)}{\Upsilon u_L^{CD} k(\theta_L)}\right]^{\tau} = \frac{x_H}{x_L}, \tag{A2.11}$$

where we denote

$$k(\theta) = \left[\theta + (1-\theta)\lambda^{1-\gamma}\right]^{\frac{1}{1-\gamma}}, \tag{A2.12}$$

$\tau = \frac{\beta(\gamma-1)}{1+(1-\beta)(\gamma-1)}$*, and Υ is defined in (A2.9).*

Proof. By Proposition 1, the optimal risk sharing condition (23) must hold with equality for the marginal worker with the realization of $\underline{\varepsilon}_L(u_0)$ at node L. Comparing the optimal risk-sharing conditions for consumption at node H and at L, we have

$$\left[\frac{c_H(u_0)}{e^{\underline{\varepsilon}_L(u_0)} c_L(u_0,\ \underline{\varepsilon}_L(u_0))}\right]^{-1} \left[\frac{u_H(u_0)}{e^{\underline{\varepsilon}_L(u_0)} u_L(u_0,\ \underline{\varepsilon}_L(u_0))}\right]^{1-\gamma} = \left[\frac{x_H}{x_L}\right]^{-1} \left[\frac{w_H}{w_L}\right]^{1-\gamma}. \tag{A2.13}$$

We can use the promise-keeping constraint to represent continuation utilities as functions of consumption. For capital owners,

$$\begin{aligned} w_H &= x_H^{1-\beta} n_H^{\beta}, \quad \text{where } n_H = (1-\alpha) e^{g_H} u_H^{FB}, \\ w_L &= x_L^{1-\beta} n_L^{\beta}, \quad \text{where } n_L = (1-\alpha) e^{g_L} u_L^{FB}, \end{aligned} \tag{A2.14}$$

where the computation of continuation utility n_H and n_L uses the fact that capital owners are not exposed to idiosyncratic risks and that together they consume $1-\alpha$ fraction of aggregate output. Because workers are not exposed to idiosyncratic risks at node H and consume α fraction of aggregate output, their continuation utility at node H can be computed using

$$u_H(u_0) = [c_H(u_0)]^{1-\beta} m_H^{\beta}, \quad \text{where } m_H = \alpha u_H^{FB} e^{g_H} k(\theta_H), \tag{A2.15}$$

where $k(\theta)$ is defined in (A2.12). At node L, workers consume αy_t for $t = 2, 3, \ldots$. In period 2, following node L, a worker stays employed with probability θ_L, in which case his output is $y_2 = y_1 e^{g_L + \varepsilon'}$. With probability $1-\theta_L$, a worker loses $1-\lambda$ fraction of human capital and his output is $y_2 = \lambda y_1 e^{g_L+\varepsilon'}$. Therefore, the certainty equivalent for a worker at node L is

$$\begin{aligned} m_L &= \left\{\int_{-\infty}^{\infty} \left[e^{g'+\varepsilon'}\left(\theta_L \alpha u_L^{CD} + (1-\theta_L)\lambda\alpha u_L^{CD}\right)\right]^{1-\gamma} f(\varepsilon' | g_L)\, d\varepsilon'\right\}^{\frac{1}{1-\gamma}} \\ &= \alpha \Upsilon u_L^{CD} e^{g_L} k(\theta_L), \end{aligned} \tag{A2.16}$$

where we define $\Upsilon \in (0,1)$ as in (A2.9). Therefore, the normalized utility of the marginal agent at node L can be written as

$$u_L(u_0, \underline{\varepsilon}_L(u_0)) = [c_L(u_0,\ \underline{\varepsilon}_L(u_0))]^{1-\beta} m_L^{\beta}, \quad \text{where } m_L = \alpha\Upsilon u_L^{CD} e^{g_L} k(\theta_L). \tag{A2.17}$$

Now we use expressions in (A2.14) and (A2.17) to replace the continuation utilities in (A2.13) and simplify to get

$$\left[\frac{c_H(u_0)}{e^{[1+\tau]\underline{\varepsilon}_L(u_0)}c_L(u_0,\ \underline{\varepsilon}_L(u_0))}\right]^{-\Omega}\left[\frac{\Upsilon u_L^{CD}k(\theta_L)}{u_L^{FB}k(\theta_H)}\right]^{-\beta(1-\gamma)}=\left[\frac{x_H}{x_L}\right]^{-\Omega}, \tag{A2.18}$$

where to simplify notation, we denote

$$\Omega\equiv 1+(1-\beta)(\gamma-1)>0,\quad \textit{and}\quad \tau\equiv\frac{\beta(\gamma-1)}{\Omega}. \tag{A2.19}$$

We can therefore obtain (A2.11) by raising both sides of equation (A2.13) to their $-\frac{1}{\Omega}th$ power. ∎

Next, we provide a lemma that links the consumption of a marginal worker to the expected consumption of an average worker at node L.

Lemma 7. *(Expected worker consumption at node L)*

Given the consumption share of the capital owners, x_H and x_L, the expected consumption of a worker with promised utility u_0 at node L is given by:

$$E\left[e^{\varepsilon'}c_L(u_0,\varepsilon')\right]=e^{(1+\tau)\underline{\varepsilon}_L(u_0)}c_L(u_0,\underline{\varepsilon}_L(u_0))\,\Phi(\underline{\varepsilon}_L(u_0)), \tag{A2.20}$$

for all u_0, where the function $\Phi(\varepsilon)$ is defined as

$$\Phi(\varepsilon)=\frac{\xi}{\xi-\tau}e^{-\tau\varepsilon_{MAX}}-\frac{\xi(1+\tau)}{(1+\xi)(\xi-\tau)}e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\varepsilon}. \tag{A2.21}$$

Proof. Note that $\forall\varepsilon'\leq\underline{\varepsilon}_L(u_0)$, the limited commitment constraint binds, and $c_L(u_0,\ \varepsilon')=c_L(u_0,\ \underline{\varepsilon}_L(u_0))$. Therefore, the expected consumption of a worker with promised utility u_0 at node L can be computed as

$$\int_{-\infty}^{\underline{\varepsilon}_L(u_0)}e^{\varepsilon'}c_L(u_0,\ \underline{\varepsilon}_L(u_0))\,f(\varepsilon'|\,g_L)\,d\varepsilon'+\int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}}e^{\varepsilon'}c_L(u_0,\ \varepsilon')\,f(\varepsilon'|\,g_L)\,d\varepsilon'. \tag{A2.22}$$

To compute $c_L(u_0,\ \varepsilon')$, note that the first order condition (23) implies that for all $\varepsilon'\geq\underline{\varepsilon}_L(u_0)$,

$$e^{-\gamma\varepsilon'}[c_L(u_0,\ \varepsilon')]^{-1}[u_L(u_0,\varepsilon')]^{1-\gamma}=e^{-\gamma\underline{\varepsilon}_L(u_0)}[c_L(u_0,\ \underline{\varepsilon}_L(u_0))]^{-1}[u_L(u_0,\ \underline{\varepsilon}_L(u_0))]^{1-\gamma}. \tag{A2.23}$$

We can compute $u_L(u_0,\varepsilon')$ as:

$$u_L(u_0,\varepsilon')=c_L^{1-\beta}(u_0,\ \varepsilon')\,m_L^{\beta}, \tag{A2.24}$$

where the expression of m_L is given in equation (A2.16). Equations (A2.23) and (A2.24) together imply

$$e^{-\gamma\varepsilon'}[c_L(u_0,\varepsilon')]^{-1+(1-\gamma)(1-\beta)}=e^{-\gamma\underline{\varepsilon}_L(u_0)}[c_L(u_0,\ \underline{\varepsilon}_L(u_0))]^{-1+(1-\gamma)(1-\beta)}.$$

Raising both sides of the above equation to the $-\frac{1}{\Omega}th$ power and using the definition of Ω and τ in (A2.19), we have, for all $\varepsilon\geq\underline{\varepsilon}_L(u_0)$,

$$e^{\varepsilon'}c_L(u_0,\varepsilon')=e^{-\tau\varepsilon'}e^{(1+\tau)\underline{\varepsilon}_L(u_0)}c_L(u_0,\ \underline{\varepsilon}_L(u_0)). \tag{A2.25}$$

Now, we compute the first term in the integral in (A2.22) as:

$$\int_{-\infty}^{\underline{\varepsilon}_L(u_0)} e^{\varepsilon'} c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right) f\left(\varepsilon' \middle| g_L\right) d\varepsilon' = c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right) \int_{-\infty}^{\underline{\varepsilon}_L(u_0)} e^{\varepsilon'} f\left(\varepsilon' \middle| g_L\right) d\varepsilon' = \frac{\xi}{1+\xi} e^{-\xi\varepsilon_{MAX}+(1+\xi)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right), \tag{A2.26}$$

and the second term as

$$\begin{aligned}
&\int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}} e^{\varepsilon'} c_L\left(u_0,\ \varepsilon'\right) f\left(\varepsilon' \middle| g_L\right) d\varepsilon' \\
=\ & e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right) \int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}} e^{-\tau\varepsilon'} f\left(\varepsilon' \middle| g_L\right) d\varepsilon \\
=\ & e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right) \frac{\xi}{\xi-\tau} \left[e^{-\tau\varepsilon_{MAX}} - e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\underline{\varepsilon}_L(u_0)}\right]
\end{aligned} \tag{A2.27}$$

We obtain equation (A2.20) by summing up (A2.26) and (A2.27). ∎

Lemma 6 is the optimal risk sharing condition that equalizes the marginal rate of substitution of workers and capital owners across the two states in period 1. The next lemma provides another first-order condition that links the marginal rate of substitution of capital owners and workers across time. Together Lemma 6 and Lemma 8 below completely characterize optimal risk sharing conditions.

Lemma 8. *(Optimal risk sharing)*

Optimal risk sharing requires that for all u_0,

$$\left[\frac{x_H}{c_H(u_0)}\right]^{1+(1-\beta)(\gamma-1)} = \left[\frac{x_0}{c_0(u_0)}\right]\left[\frac{\bar{n}_0(x_H, x_L)}{\bar{m}_0(u_0)}\right]^{\gamma-1}, \tag{A2.28}$$

where

$$\bar{n}_0(x_H, x_L) = \left[\pi\left(e^{(1+\beta)g_H} x_H^{(1-\beta)}\left(u_H^{FB}\right)^{\beta}\right)^{1-\gamma} + (1-\pi)\left(e^{(1-\gamma)g_L} x_L^{(1-\beta)}\left(u_L^{FB}\right)^{\beta}\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}},$$

and

$$\bar{m}_0(u_0) = \left[\begin{array}{c} \pi\left(e^{(1+\beta)g_H} c_H^{1-\beta}\left(u_H^{FB} k(\theta_H)\right)^{\beta}\right)^{1-\gamma} + \\ (1-\pi) e^{(1-\gamma)g_L}\left[e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L(u_0)\right)\right]^{(1-\beta)(1-\gamma)}\left\{\frac{1}{\alpha}m_L\right\}^{\beta(1-\gamma)}\Psi\left(\underline{\varepsilon}_L(u_0)\right)\end{array}\right]^{\frac{1}{1-\gamma}}, \tag{A2.29}$$

where $\Psi(\varepsilon)$ is given by:

$$\Psi(\varepsilon) = \left\{\frac{\xi}{\xi-\tau} e^{-\tau\varepsilon_{MAX}} - \frac{\xi(1-\gamma+\tau)}{(\xi-\tau)(\xi+1-\gamma)} e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\underline{\varepsilon}_L(u_0)}\right\}. \tag{A2.30}$$

Proof. The optimal risk sharing condition implies that

$$\left[\frac{c_H(u_0)}{c_0}\right]^{-1}\left[\frac{u_H(u_0)}{m_0}\right]^{1-\gamma} = \left[\frac{x_H}{x_0}\right]^{-1}\left[\frac{w_H}{n_0}\right]^{1-\gamma}. \tag{A2.31}$$

Using equation (A2.14),

$$\begin{aligned} w_H &= x_H^{1-\beta}\left[(1-\alpha)\,e^{g_H}u_H^{FB}\right]^{\beta},\ w_L = x_L^{1-\beta}\left[(1-\alpha)\,e^{g_L}u_L^{FB}\right]^{\beta},\ \ \textit{and} \\ n_0 &= \left[\pi\left(e^{g_H}w_H\right)^{1-\gamma}+(1-\pi)\left(e^{g_L}w_L\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}}. \end{aligned} \tag{A2.32}$$

To calculate workers' utility, use (A2.15) to obtain

$$u_H(u_0) = \left[c_H(u_0)\right]^{1-\beta}\left[\alpha u_H^{FB}e^{g_H}k(\theta_H)\right]^{\beta}. \tag{A2.33}$$

Using equations (A2.32) and (A2.33) to replace the relevant terms in (A2.31), we obtain equation (A2.28). It remains to calculate workers' certainty equivalent,

$$m_0 = \left\{\pi\left[e^{g_H}u_H(u_0)\right]^{1-\gamma}+(1-\pi)\int_{-\infty}^{\infty}\left[e^{g_L+\varepsilon'}u_L(u_0,\varepsilon')\right]^{1-\gamma}f(\varepsilon'|\,g_L)\,d\varepsilon'\right\}^{\frac{1}{1-\gamma}}. \tag{A2.34}$$

Note that for $\varepsilon' \geq \underline{\varepsilon}_L(u_0)$, using equation (A2.23), we can write

$$\begin{aligned} \left[e^{\varepsilon'}u_L(u_0,\varepsilon')\right]^{1-\gamma} &= \left[e^{\varepsilon'}c_L(u_0,\ \varepsilon')\right]e^{-\gamma\underline{\varepsilon}_L(u_0)}\left[c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{-1}\left[u_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{1-\gamma} \\ &= \left[e^{\varepsilon'}c_L(u_0,\ \varepsilon')\right]e^{-\gamma\underline{\varepsilon}_L(u_0)}\left[c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{-1+(1-\beta)(1-\gamma)}m_L^{\beta(1-\gamma)}, \end{aligned}$$

where the second equality uses (A2.17) to compute $u_L(u_0,\ \underline{\varepsilon}_L(u_0))$ as a function of consumption. Therefore,

$$\begin{aligned} &\int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}}\left[e^{\varepsilon'}u_L(u_0,\varepsilon')\right]^{1-\gamma}f(\varepsilon'|\,g_L)\,d\varepsilon' \\ =\ &\int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}}\left[e^{\varepsilon'}c_L(u_0,\ \varepsilon')\right]f(\varepsilon'|\,g_L)\,d\varepsilon'\times e^{-\gamma\underline{\varepsilon}_L(u_0)}\left[c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{-1+(1-\beta)(1-\gamma)}m_L^{\beta(1-\gamma)} \\ =\ &\frac{\xi}{\xi-\tau}\left[e^{-\tau\varepsilon_{MAX}}-e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\underline{\varepsilon}_L(u_0)}\right]\times e^{(1-\gamma+\tau)\underline{\varepsilon}_L(u_0)}\left[c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{(1-\beta)(1-\gamma)}m_L^{\beta(1-\gamma)} \\ =\ &\frac{\xi}{\xi-\tau}\left[e^{-\tau\varepsilon_{MAX}}-e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\underline{\varepsilon}_L(u_0)}\right]\left[e^{(1+\tau)\underline{\varepsilon}_L(u_0)}c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{(1-\beta)(1-\gamma)}m_L^{\beta(1-\gamma)}, \end{aligned} \tag{A2.35}$$

where the second equality uses the same calculation as in (A2.27) and the last equality uses the definition of τ to simplify. For $\varepsilon' < \underline{\varepsilon}_L(u_0)$, the firm-side limited commitment constraint binds, and

$$\begin{aligned} \left[e^{\varepsilon'}u_L(u_0,\varepsilon')\right]^{1-\gamma} &= e^{(1-\gamma)\varepsilon'}\left[u_L(u_0,\underline{\varepsilon}_L(u_0))\right]^{1-\gamma} \\ &= e^{(1-\gamma)\varepsilon'}\left[c_L(u_0,\ \underline{\varepsilon}_L(u_0))\right]^{(1-\beta)(1-\gamma)}m_L^{\beta(1-\gamma)}, \end{aligned}$$

where the second equality applies equation (A2.17). Therefore,

$$\begin{aligned}
& \int_{-\infty}^{\underline{\varepsilon}_L(u_0)} \left[e^{\varepsilon'} u_L\left(u_0, \varepsilon'\right)\right]^{1-\gamma} f\left(\varepsilon' \middle| g_L\right) d\varepsilon' \\
= & \int_{-\infty}^{\underline{\varepsilon}_L(u_0)} e^{(1-\gamma)\varepsilon'} f\left(\varepsilon' \middle| g_L\right) d\varepsilon' \times \left[c_L\left(u_0,\ \underline{\varepsilon}_L\left(u_0\right)\right)\right]^{(1-\beta)(1-\gamma)} m_L^{\beta(1-\gamma)} \\
= & \frac{\xi}{\xi+1-\gamma} e^{-\xi\varepsilon_{MAX}+(1-\gamma+\xi)\underline{\varepsilon}_L(u_0)} \times \left[c_L\left(u_0,\ \underline{\varepsilon}_L\left(u_0\right)\right)\right]^{(1-\beta)(1-\gamma)} m_L^{\beta(1-\gamma)} \\
= & \frac{\xi}{\xi+1-\gamma} e^{-\xi\varepsilon_{MAX}+(\xi-\tau)\underline{\varepsilon}_L(u_0)} \times \left[e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L\left(u_0\right)\right)\right]^{(1-\beta)(1-\gamma)} m_L^{\beta(1-\gamma)}, \quad \text{(A2.36)}
\end{aligned}$$

where the last equality uses the definition of τ to simplify. Combining (A2.35) and (A2.36), we have

$$\begin{aligned}
& \int_{-\infty}^{\infty} \left[e^{\varepsilon'} u_L\left(u_0, \varepsilon'\right)\right]^{1-\gamma} f\left(\varepsilon' \middle| g_L\right) d\varepsilon' \\
= & \left[e^{(1+\tau)\underline{\varepsilon}_L(u_0)} c_L\left(u_0,\ \underline{\varepsilon}_L\left(u_0\right)\right)\right]^{(1-\beta)(1-\gamma)} m_L^{\beta(1-\gamma)} \Psi\left(\underline{\varepsilon}_L\left(u_0\right)\right), \quad \text{(A2.37)}
\end{aligned}$$

where $\underline{\varepsilon}_L\left(u_0\right)$ is defined in (A2.30). We obtain the expression (A2.29) by combing (A2.34) and (A2.37). ∎

General Equilibrium Unit measure of a single type of workers and market clearing at node 0, node H, and node L implies u_0^* solves

$$c_0\left(u_0^*\right) = 1 - x_0,\ c_H\left(u_0^*\right) = 1 - x_H, \ \ \textit{and}\ E\left[e^{\varepsilon'} c_L\left(u_0^*, \varepsilon'\right)\right] = 1 - x_L,$$

respectively. Note that equation (A2.10) implies

$$c_L\left(u_0^*, \underline{\varepsilon}_L\left(u_0^*\right)\right) = 1 + \frac{\beta}{1-\beta} x_L - a \ln\left[1 + \frac{\beta x_L}{a\left(1-\beta\right)}\right]. \quad \text{(A2.38)}$$

Using market clearing at node L and Lemma 7,

$$1 - x_L = e^{(1+\tau)\underline{\varepsilon}_L(u_0^*)} c_L\left(u_0, \underline{\varepsilon}_L\left(u_0^*\right)\right) \Phi\left(\underline{\varepsilon}_L\left(u_0^*\right)\right). \quad \text{(A2.39)}$$

Combining (A2.38) and (A2.39), we have:

$$e^{(1+\tau)\underline{\varepsilon}_L(u_0^*)} \Phi\left(\underline{\varepsilon}_L\left(u_0^*\right)\right) = \frac{1 - x_L}{1 + \frac{\beta}{1-\beta} x_L - a \ln\left[1 + \frac{\beta x_L}{a(1-\beta)}\right]}. \quad \text{(A2.40)}$$

Equations (A2.38) and (A2.40) together define $c_L\left(u_0, \underline{\varepsilon}_L\left(u_0^*\right)\right)$ and $\underline{\varepsilon}_L\left(u_0^*\right)$ as functions of x_L. With a bit abuse of notation, we denote these functions as $c_L\left(x_L\right)$ and $\underline{\varepsilon}(x_L)$.

Focusing on type-u_0^* agents, using Lemma 7, we can replace the term $e^{(1+\tau)\underline{\varepsilon}_L(u_0^*)} c_L\left(u_0^*, \underline{\varepsilon}_L\left(u_0^*\right)\right)$ in equation (A2.11) by the following

$$e^{(1+\tau)\underline{\varepsilon}_L(u_0^*)} c_L\left(u_0^*, \underline{\varepsilon}_L\left(u_0^*\right)\right) = \left(1 - x_L\right) \Phi\left(\underline{\varepsilon}\left(x_L\right)\right)^{-1}. \quad \text{(A2.41)}$$

Therefore, the first order condition (A2.11) can be written as

$$\Phi\left(\underline{\varepsilon}\left(x_L\right)\right) \left[\frac{u_L^{FB} k\left(\theta_H\right)}{\Upsilon u_L^{CD} k\left(\theta_L\right)}\right]^{\tau} = \frac{x_H}{x_L} \frac{1 - x_L}{1 - x_H}. \quad \text{(A2.42)}$$

Also, using the marketing clearing condition to replace c_H by $1 - x_H$, and use (A2.41) to replace $e^{(1+\tau)\underline{\varepsilon}_L(u_0^*)} c_L\left(u_0^*, \underline{\varepsilon}_L\left(u_0^*\right)\right)$, we define workers' certainty equivalent as a function of x_H, x_L, and ε using (A2.29)

$$\bar{m}_0\left(x_H, x_L, \varepsilon\right) = \left\{ \begin{array}{c} \pi\left[e^{(1+\beta)g_H}\left(1-x_H\right)^{(1-\beta)}\left(u_H^{FB} k\left(\theta_H\right)\right)^{\beta}\right]^{1-\gamma} \\ +\left(1-\pi\right)\left[e^{(1+\beta)g_L}\left[\frac{1-x_L}{\Phi(\varepsilon)}\right]^{(1-\beta)}\left[\Upsilon u_L^{CD} k\left(\theta_L\right)\right]^{\beta}\right]^{1-\gamma} \Psi\left(\varepsilon\right) \end{array} \right\}^{\frac{1}{1-\gamma}}, \tag{A2.43}$$

and the first order condition (A2.28) can be written as

$$\left[\frac{x_H}{1-x_H}\right]^{1+(1-\beta)(\gamma-1)} = \left[\frac{x_0}{1-x_0}\right]^{-1}\left[\frac{\bar{n}_0\left(x_H, x_L\right)}{\bar{m}_0\left(x_H, x_L, \underline{\varepsilon}\left(x_L\right)\right)}\right]^{1-\gamma}. \tag{A2.44}$$

Give an inital condition of x_0, equations (A2.42) and (A2.44) can be jointed solved for x_H and x_L. Other equilibrium quantities can then be constructed analogously.

A2.2 Proof of Proposition 3

1. From the definition of u_L^{CD} in (A2.3), it is clear that as $\gamma \to 1+\xi$, $u_L^{CD} \to 0$. Consider equation (A2.42). It is straightforward to verify that $\Phi\left(\varepsilon\right)$ is strictly positive and bounded (see equation (A2.21)). Also, both $k\left(\theta_H\right)$ and $k\left(\theta_L\right)$ are bounded. Therefore, as $\gamma \to 1+\xi$ the left hand side converges to ∞, and we must have $\frac{x_H}{x_L} \to \infty$. By continuity, there exists $\hat{\gamma} \in (1, 1+\xi)$ such that $\frac{x_H}{x_L} > 1$ if and only if $\gamma > \hat{\gamma}$, as needed.

   In addition, if $\gamma = 1$, then $\tau = 0$. Using the definition of $\Phi\left(\varepsilon\right)$,

   $$\Phi\left(\varepsilon\right) = 1 - \frac{1}{\left(1+\xi\right)} e^{-\xi\left(\varepsilon_{MAX} - \varepsilon\right)} < 1.$$

   Therefore, we must have $\frac{x_H}{x_L} < 1$.

2. The economy without moral hazard is a special case in which the parameter for cost of effort, $a = 0$. We use $\theta_H\left(a\right)$ and $\theta_L\left(a\right)$ to denote policy functions with the understanding that they are policy functions of the moral hazard economy if $a > 0$, and they stand for policy functions in the economy without moral hazard if $a = 0$. Using our result from Part 1 of the proof, as $\gamma \to 1+\xi$, $\frac{x_H}{x_L} \to \infty$. Because both x_H and x_L are bounded between $[0, 1]$, we must have $x_L \to 0$. Therefore, $\theta_L\left(a\right) \to 0$ by equation (A2.6). Also, equation (A2.44) implies that as $\gamma \to 1+\xi$, $\bar{m}_0\left(x_H\left(a\right), x_L\left(a\right), \underline{\varepsilon}\left(x_L\left(a\right)\right)\right) \to 0$; therefore, $x_H\left(a\right) \to 0$ as well. Therefore, as $\gamma \to 1+\xi$, $\theta_H\left(a\right) \to 1 - \frac{a}{a + \frac{\beta}{1-\beta} x_H^*}$. Consider equation (A2.42), for an arbitrary a,

   $$\left[\frac{k\left(\theta_H\left(a\right)\right)}{k\left(\theta_L\left(a\right)\right)}\right]^{\tau} = \left[\frac{\theta_H\left(a\right) + \left(1-\theta_H\left(a\right)\right)\lambda^{1-\gamma}}{\theta_L\left(a\right) + \left(1-\theta_L\left(a\right)\right)\lambda^{1-\gamma}}\right]^{\frac{\tau}{1-\gamma}}.$$

   Suppose $a > 0$, then as $\gamma \to 1+\xi$, there exist $\epsilon > 0$ such that

   $$\left[\frac{\theta_H\left(a\right) + \left(1-\theta_H\left(a\right)\right)\lambda^{1-\gamma}}{\theta_L\left(a\right) + \left(1-\theta_L\left(a\right)\right)\lambda^{1-\gamma}}\right]^{\frac{\tau}{1-\gamma}} \to \left[\frac{1 - \frac{a}{a + \frac{\beta}{1-\beta} x_H^*(a)} + \left(1 - \frac{a}{a + \frac{\beta}{1-\beta} x_H^*(a)}\right)\lambda^{-\xi}}{\lambda^{-\xi}}\right]^{-\frac{1}{\xi}\frac{\beta\xi}{1+\xi(1-\beta)}} > 1+\epsilon.$$

   In addition, equation (A2.40) implies that as $\gamma \to 1+\xi$, $x_L \to 0$, and therefore, $\underline{\varepsilon}_L\left(a\right) \to \varepsilon^*$

for all a, where ε^* is such that $e^{(1+\tau)\varepsilon^*}\Phi(\varepsilon^*) = 1$. Therefore, with $a > 0$, for γ close enough to $1+\xi$, we must have

$$\Phi\left(\underline{\varepsilon}_L(a)\right)\left[\frac{u_L^{FB}k\left(\theta_H(a)\right)}{\Upsilon u_L^{CD}k\left(\theta_L(a)\right)}\right]^\tau > \Phi\left(\underline{\varepsilon}_L(0)\right)\left[\frac{u_L^{FB}}{\Upsilon u_L^{CD}}\right]^\tau.$$

Equation (A2.42) implies that for γ close enough to $1+\xi$, $\frac{x_H}{x_L} > \frac{x_H}{x_L}$ because as $\gamma \to 1+\xi$, $x_L \to 0$ and $x_H \to x_H^*$ has a limit.

3. By Part 1 of the proposition, for γ large enough, $x_H > x_L$. The fact that $\theta_H > \theta_L$ follows from equation (A2.6).

Proof for the claim that Price-dividend ratio is procyclical Here we provide a proof for claim in footnote 13. Consider first firm value at node H, (A2.4). Because there is no idiosyncratic shock at node H, there is only one type of firm, and $u = u_H$. Using the market clearing condition at node H, $1 - c_H = x_H$. Therefore, in equilibrium,

$$\begin{aligned}
v_H\left(u_0^*\right) &= 1 - c_H\left(u_0^*\right) + \frac{\beta}{1-\beta}x_H - a\ln\left[1 + \frac{\beta x_H}{a(1-\beta)}\right] \\
&= x_H + \frac{\beta}{1-\beta}x_H - a\ln\left[1 + \frac{\beta x_H}{a(1-\beta)}\right] \\
&= \frac{1}{1-\beta}x_H - a\ln\left[1 + \frac{\beta x_H}{a(1-\beta)}\right].
\end{aligned}$$

At node L, firm value is given by $e^\varepsilon v_L\left(u_0^*, \varepsilon\right)$. Using equation (A2.5),

$$v_L\left(u_0^*, \varepsilon\right) = 1 - c_L\left(u_0^*, \varepsilon\right) + \frac{\beta}{1-\beta}x_L - a\ln\left[1 + \frac{\beta x_L}{a(1-\beta)}\right].$$

Note that

$$E\left[e^\varepsilon c_L\left(u_0^*, \varepsilon\right)\right] = 1 - x_L$$

by market clearing. Therefore,

$$\begin{aligned}
E\left[e^\varepsilon v_L\left(u_0^*, \varepsilon\right)\right] &= 1 - E\left[e^\varepsilon c_L\left(u_0^*, \varepsilon\right)\right] + \frac{\beta}{1-\beta}x_L - a\ln\left[1 + \frac{\beta x_L}{a(1-\beta)}\right] \\
&= \frac{1}{1-\beta}x_L - a\ln\left[1 + \frac{\beta x_L}{a(1-\beta)}\right].
\end{aligned}$$

Using our previous arguement, as $\gamma \to 1+\xi$, $\frac{x_H}{x_L} \to \infty$. Therefore, for γ large enough, we must have $\frac{v_H\left(u_0^*\right)}{E\left[e^\varepsilon v_L\left(u_0^*, \varepsilon\right)\right]} > 1$, as needed.

A2.3 Proof of Proposition 4

Firm risk pass through Fixing u_0, equation (A2.25) implies that $\forall \varepsilon' \geq \underline{\varepsilon}(u_0)$,

$$\frac{d\ln\left[e^\varepsilon c_L\left(u_0, \varepsilon\right)\right]}{d\varepsilon} = -\tau.$$

For $\varepsilon' < \underline{\varepsilon}(u_0)$, the limited commitment constraint binds, and $e^{\varepsilon}c_L(u_0,\varepsilon) = e^{\varepsilon}c_L(u_0,\underline{\varepsilon}(u_0))$. Therefore,

$$\frac{d\ln\left[e^{\varepsilon}c_L(u_0,\varepsilon)\right]}{d\varepsilon} = 1$$

Combining the above two equations, we have

$$\begin{aligned} E\left[\frac{\partial \ln\left[e^{\varepsilon}c_L(u_0,\varepsilon)\right]}{\partial\varepsilon}\right] &= \int_{-\infty}^{\underline{\varepsilon}_L(u_0)} f(\varepsilon'|\, g_L)\, d\varepsilon'\tau + \int_{\underline{\varepsilon}_L(u_0)}^{\varepsilon_{MAX}} f(\varepsilon'|\, g_L)\, d\varepsilon' \\ &= e^{-\xi\left(\varepsilon_{MAX}-\underline{\varepsilon}_L(u_0)\right)} - \tau\left[1-e^{-\xi\left(\varepsilon_{MAX}-\underline{\varepsilon}_L(u_0)\right)}\right]. \end{aligned}$$

Clearly, the average elasticity is increasing in $\underline{\varepsilon}_L(u_0)$. Using the optimal risk sharing conditions (A2.13) and (A2.28), we can show that $\underline{\varepsilon}_L(u_0)$ is an increasing function of u_0.

Cross section of expected returns To characterize the dependence of $\frac{v_H(u_0)}{E[e^{\varepsilon}v_L(u_0,\varepsilon)]}$, note that in general,

$$c_H(u_0) = \frac{x_H}{x_L}\left[\frac{\xi u_L^{CD}k(\theta_L)}{u_L^{FB}k(\theta_L)}\right]^{\tau} e^{(1+\tau)\underline{\varepsilon}_L(u_0)}c_L\left(u_0,\underline{\varepsilon}_L(u_0)\right)$$

by Lemma 6 and

$$E\left[e^{\varepsilon}c_L(u_0,\varepsilon)\right] = e^{(1+\tau)\underline{\varepsilon}_L(u_0)}c_L(u_0,\underline{\varepsilon}_L(u_0))\,\Phi(\underline{\varepsilon}_L(u_0))$$

by Lemma 7. Because at $\varepsilon = \underline{\varepsilon}_L(u_0)$, the limited commitment constraint, $v_L(u_0,\varepsilon) = 0$ binds, $c_L(u_0,\underline{\varepsilon}_L(u_0)) = 1 + \frac{\beta}{1-\beta}x_L - a\ln\left[1+\frac{\beta x_L}{a(1-\beta)}\right]$ by (A2.10). To simplify notation, we denote $\Gamma_H = 1+\frac{\beta}{1-\beta}x_H - a\ln\left[1+\frac{\beta x_H}{a(1-\beta)}\right]$ and $\Gamma_L = 1+\frac{\beta}{1-\beta}x_L - a\ln\left[1+\frac{\beta x_L}{a(1-\beta)}\right]$. We then write $\frac{v_H(u_0)}{E[e^{\varepsilon}v_L(u_0,\varepsilon)]}$ as

$$\frac{v_H(u_0)}{E\left[e^{\varepsilon}v_L(u_0,\varepsilon)\right]} = \frac{\Gamma_H - \phi e^{(1+\tau)\underline{\varepsilon}_L(u_0)}}{\Gamma_L\left\{1-e^{(1+\tau)\underline{\varepsilon}_L(u_0)}\Phi\left(\underline{\varepsilon}(u_0)\right)\right\}},$$

where we denote $\phi = \frac{x_H}{x_L}\left[\frac{\xi u_L^{CD}k(\theta_L)}{u_L^{FB}k(\theta_L)}\right]^{\tau}\Gamma_L$ to simplify notation. By Proposition 1, $\underline{\varepsilon}(u_0)$ is a strictly increasing function of u_0. Therefore, to prove Proposition 4, it enough to show

$$\frac{\partial}{\partial\underline{\varepsilon}}\frac{\Gamma_H-\phi e^{(1+\tau)\underline{\varepsilon}}}{\left\{1-e^{(1+\tau)\underline{\varepsilon}}\Phi\left(\underline{\varepsilon}\right)\right\}} > 0,$$

which is given by the following lemma.

Lemma 9. *There exists $\tilde{\gamma} \in (1, 1+\xi)$ such that $\gamma > \tilde{\gamma}$ implies that for all $\varepsilon \in (-\infty, \varepsilon_{MAX})$,*

$$\frac{\partial}{\partial\varepsilon}\left[\frac{\Gamma_H - \phi e^{(1+\tau)\varepsilon}}{1-e^{(1+\tau)\varepsilon}\Phi(\varepsilon)}\right] > 0. \tag{A2.45}$$

Proof. We can compute (A2.45) as :

$$\begin{aligned} &\frac{\partial}{\partial\varepsilon}\left[\frac{\Gamma_H-\phi e^{(1+\tau)\varepsilon}}{1-e^{(1+\tau)\varepsilon}\Phi(\varepsilon)}\right] \\ = \;& \frac{-\phi e^{(1+\tau)\varepsilon}(1+\tau)\left[1-e^{(1+\tau)\varepsilon}\Phi(\varepsilon)\right] + \left[\Gamma_H-\phi e^{(1+\tau)\varepsilon}\right]e^{(1+\tau)\varepsilon}\left[(1+\tau)\Phi(\varepsilon)+\Phi'(\varepsilon)\right]}{\left[1-e^{(1+\tau)\varepsilon}\Phi(\varepsilon)\right]^2}. \end{aligned}$$

We focus on the numerator and simplify:

$$\begin{aligned} & -\phi e^{(1+\tau)\varepsilon}(1+\tau)\left[1-e^{(1+\tau)\varepsilon}\Phi(\varepsilon)\right]+\left[\Gamma_H-\phi e^{(1+\tau)\varepsilon}\right]e^{(1+\tau)\varepsilon}\left[(1+\tau)\Phi(\varepsilon)+\Phi'(\varepsilon)\right] \\ = & \ \Gamma_H\left[(1+\tau)\Phi(\varepsilon)+\Phi'(\varepsilon)\right]-\phi\left[(1+\tau)+e^{(1+\tau)\varepsilon}\Phi(\varepsilon)\right] \end{aligned}$$

It is therefore enough to show

$$\Gamma_H\left[(1+\tau)\Phi(\varepsilon)+\Phi'(\varepsilon)\right]-\phi\left[(1+\tau)+e^{(1+\tau)\varepsilon}\Phi(\varepsilon)\right]>0 \tag{A2.46}$$

Using the expression of $\Phi(\varepsilon)$, we can compute

$$\begin{aligned} (1+\tau)\Phi(\varepsilon)+\Phi'(\varepsilon) &= (1+\tau)\frac{\xi}{\xi-\tau}\left[e^{-\tau\varepsilon_{MAX}}-e^{-\lambda\varepsilon_{MAX}+(\xi-\tau)\varepsilon}\right] \\ &= (1+\tau)\frac{\xi}{\xi-\tau}e^{-\tau\varepsilon_{MAX}}\left[1-e^{-(\xi-\tau)\varepsilon_{MAX}+(\xi-\tau)\varepsilon}\right] \\ &= (1+\tau)\frac{\xi}{1+\xi}e^{-\tau\varepsilon_{MAX}}\frac{1+\xi}{\xi-\tau}\left[1-e^{-(\xi-\tau)\varepsilon_{MAX}+(\xi-\tau)\varepsilon}\right] \\ &= (1+\tau)e^{-(1+\tau)\varepsilon_{MAX}}\frac{1+\xi}{\xi-\tau}\left[1-e^{-(\xi-\tau)(\varepsilon_{MAX}-\varepsilon)}\right]>0, \end{aligned}$$

where the last line uses the fact $\varepsilon_{MAX}=\ln\frac{1+\xi}{\xi}$. Also, the second term in equation (A2.46) can be written as

$$\begin{aligned} (1+\tau)+e^{(1+\tau)\varepsilon}\Phi'(\varepsilon) &= (1+\tau)\left[1-\frac{\xi}{1+\xi}e^{-\lambda\varepsilon_{MAX}+(1+\xi)\varepsilon}\right] \\ &= (1+\tau)\left[1-e^{-(1+\xi)(\varepsilon_{MAX}-\varepsilon)}\right]. \end{aligned}$$

Therefore, to prove (A2.46), it is enough to show that for all ε,

$$\Gamma_H e^{-(1+\tau)\varepsilon_{MAX}}\frac{1+\xi}{\xi-\tau}\left[1-e^{-(\xi-\tau)(\varepsilon_{MAX}-\varepsilon)}\right]-\phi\left[1-e^{-(1+\xi)(\varepsilon_{MAX}-\varepsilon)}\right]>0.$$

Because $\phi\to 0$ as $\gamma\to 1+\xi$, we have $\Gamma_H e^{-(1+\tau)\varepsilon_{MAX}}>\phi$ for γ close enough to $1+\xi$. We complete the proof by making the following observation

Define

$$\begin{aligned} f(\varepsilon) &= \frac{1+\xi}{\xi-\tau}\left[1-e^{-(\xi-\tau)(\varepsilon_{MAX}-\varepsilon)}\right] \\ g(\varepsilon) &= 1-e^{-(1+\xi)(\varepsilon_{MAX}-\varepsilon)}, \end{aligned}$$

then $f(\varepsilon)>g(\varepsilon)$ for all $\varepsilon<\varepsilon_{MAX}$. To see this, note that $f(\varepsilon_{MAX})=g(\varepsilon_{MAX})=0$. Also, $f'(\varepsilon)<g'(\varepsilon)$ for all $\varepsilon<\varepsilon_{MAX}$, because

$$\begin{aligned} f'(\varepsilon) &= -(1+\xi)e^{-(\xi-\tau)(\varepsilon_{MAX}-\varepsilon)} \\ g'(\varepsilon) &= -(1+\xi)e^{-(1+\xi)(\varepsilon_{MAX}-\varepsilon)}. \end{aligned}$$

∎

A3 Computational Algorithm

We describe our computation algorithm. The algorithm consists of an "outer loop", in which we iterate over the law of motion for aggregate states and an associated stochastic discount factor, and an "inner loop", in which we solve for the optimal contract.

1. Initialize the law of motion of x, $\Gamma_x(g'|g,x)$. We use a log-linear functional form:

$$\log x' = a(g,g') + b(g,g')\log x. \tag{A3.1}$$

   Given the law of motion of x, the SDF $\Lambda\left(g'|\,g,x\right)$ is calculated using

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

   where $w(g,x)$ and $n(g,x)$ are given by

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

$$n\left(g,x\right) = \left\{\kappa\sum_{g'}\pi\left(g'|g\right)e^{(1-\gamma)g'}w^{1-\gamma}\left(g',\Gamma_x(g'|g,x)\right)\right\}^{\frac{1}{1-\gamma}}.$$

2. The inner loop consists of using $\Gamma_x(g'|g,x)$ and $\Lambda(g'|g,x)$, to solve the value function $v\left(u|\,g,x\right)$, the worker-outside value $\overline{u}\left(g,x\right)$ and value of a new job $u^*(g,x)$ along with the policy functions $c\left(u|\,g,x\right)$, $\theta(u|g,x)$ and $u'\left(u,s'|\,g,x\right)$ that solve the optimal contracting problem *P1*. We solve Bellman equation by a modified value function iteration as appplying a standard value function iteration is complicated by the presence of the occasionally binding constraints (12) and (13). Our procedure borrows elements from endogenous grid method of Carroll (2006). We describe it below

   (a) Guess $v(u|g,x)$ and $c(u|g,x)$. These imply functions $u^*(g,x)$ and $\overline{u}(g,x)$ using equations

$$v\left(u^*\left(g,x\right)|\,g,x\right) = 0,$$

$$\bar{u}\left(g,x\right) = \left[\left(1-\beta\right)b^{1-\frac{1}{\psi}} + \beta\left[\lambda\bar{m}\left(g,x\right)\right]^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}},$$

   with

$$\bar{m}\left(g,x\right) = \left(\kappa\mathbb{E}\left[e^{(1-\gamma)g'}\left\{(1-\chi)\bar{u}^{1-\gamma}\left(g',x(g')\right)^{1-\gamma} + \chi u^{*1-\gamma}\left(g',\Gamma_x(g'|g,x)\right)\right\}\Big|\,g\right]\right)^{\frac{1}{1-\gamma}}.$$

   We denote $c(u^*(g,x)|g,x)$ and $c(\lambda\overline{u}(g,x),g,x)$ by $c^*(g,x)$ and $\overline{c}(g,x)$.

   (b) Let $\{\underline{\varepsilon}(u,g'|g,\phi,B),\overline{\varepsilon}(u,g'|g,\phi,B)\}_{g'}$ be the thresholds for $\eta'+\varepsilon'$ such that constraint (12) and (13) bind for a worker with state u, aggregate states (ϕ,g,B) and next period for aggregate shock $g'=g_L$. Define a grid $\underline{\mathcal{E}}_L\times\mathcal{X}\equiv\left\{(\underline{\varepsilon}_{L,0},x_0),\left(\underline{\varepsilon}_{L,1},x_0\right),\ldots,\left(\underline{\varepsilon}_{L,n\mathcal{E}},x_{n\mathcal{X}}\right)\right\}$ with the understanding that $\underline{\varepsilon}_L(j)$ and $x(j)$ are the entries in the j^{th} element of the grid $\underline{\mathcal{E}}_L\times\mathcal{X}$ with $j\in\{1,2,\ldots,n\mathcal{E}\times n\mathcal{X}\}$.

   (c) For all $j\in\{1,2,\ldots,n\mathcal{E}\times n\mathcal{X}\}$, we solve for $\left\{\underline{\varepsilon}_{g'}(j),\overline{\varepsilon}_{g'}(j)\right\}_{g'}$ that are consistent with $\underline{\varepsilon}_L(j)$ and the guess for functions v and c in step (a) using the following equations that need to hold for all g'

$$\frac{\Lambda(g'|g,x(j))}{\Lambda(g_L|g,x(j))} = \frac{e^{-\gamma\left(\underline{\varepsilon}_{g'}(j)\right)}}{e^{-\gamma\left(\underline{\varepsilon}_{g_L}(j)\right)}} \left[\frac{c^*(g',\Gamma_x(g'|g,x(j)))}{c^*(g_L,\Gamma_x(g'|g,x(j)))}\right]^{-\frac{1}{\psi}} \left[\frac{u^*(g',\Gamma_x(g'|g,x(j)))}{u^*(g_L,\Gamma_x(g'|g,x(j)))}\right]^{\frac{1}{\psi}-\gamma}$$

and

$$\frac{\Lambda(g'|g,x(j))}{\Lambda(g_L|g,x(j))} = \frac{e^{-\gamma\left(\overline{\varepsilon}_{g'}(j)\right)}}{e^{-\gamma\left(\overline{\varepsilon}_{g_L}(j)\right)}} \left[\frac{\overline{c}(g',\Gamma_x(g'|g,x(j)))}{\overline{c}(g_L,\Gamma_x(g'|g,x(j)))}\right]^{-\frac{1}{\psi}} \left[\frac{\overline{u}(g',\Gamma_x(g'|g,x(j)))}{\overline{u}(g_L,\Gamma_x(g'|g,x(j)))}\right]^{\frac{1}{\psi}-\gamma}.$$

(d) Now we construct the policy function $u'(s'|j)$ using :

$$u'(\eta'+\varepsilon'|j) = u^*(g',\Gamma_x(g'|g,x(j))) \qquad \forall \eta'+\varepsilon' < \underline{\varepsilon}_{g'}(j)$$
$$u'(\eta'+\varepsilon'|j) = \lambda\overline{u}(g',\Gamma_x(g'|g,x(j))) \qquad \forall \eta'+\varepsilon' < \overline{\varepsilon}_{g'}(j)$$

and for $\eta'+\varepsilon' \in \left(\underline{\varepsilon}_{g'}(j), \overline{\varepsilon}_{g'}(j)\right)$ use

$$\frac{e^{-\gamma\left(\underline{\varepsilon}_{g'}(j)\right)}}{e^{-\gamma(\eta'+\varepsilon')}} \left[\frac{c^*(g',\Gamma_x(g'|g,x(j)))}{c(u')}\right]^{-\frac{1}{\psi}} \left[\frac{u^*(g',\Gamma_x(g'|g,x(j)))}{u'}\right]^{\frac{1}{\psi}-\gamma} = 1$$

to solve out for u'.

(e) We compute $c(j)$, $\theta(j)$ and $\iota(j)$ using

$$\Lambda(g'|g,x(j))\left(1+\frac{\iota(j)}{\theta(j)}\right) = e^{-\gamma\left(\underline{\varepsilon}_{g'}(j)\right)} \left[\frac{c^*(g',\Gamma_x(g'|g,x(j)))}{c(j)}\right]^{-\frac{1}{\psi}} \left[\frac{u^*(g',\Gamma_x(g'|g,x(j)))}{m(j)}\right]^{\frac{1}{\psi}-\gamma},$$
$$A'(\theta_j) = \kappa\mathbb{E}_g\Lambda(g'|g,x(j))\, e^{g'+\eta'+\varepsilon'} v(u'(s'|j)),$$
$$\iota(j)A''(\theta_j) = \left(\frac{\beta}{1-\beta}\right) c(j)^{\frac{1}{\psi}} m(j)^{\gamma-\frac{1}{\psi}} \mathbb{E}_g \left(\frac{1}{1-\gamma}\right) \left(e^{(1-\gamma)(\varepsilon'+\eta')}\left[u^{1-\gamma}(s'|j) - \overline{u}^{1-\gamma}(g',\Gamma_x(g'|g,x(j)))\right]\right),$$

where certainty equivalent $m(j)$ only depends on $\{u'(s'|g)\}_{s'}$ and $\{\overline{u}(g',\Gamma_x(g'|g,x(j)))\}_{g'}$.

(f) Finally, we use the promise keeping constraint (11) to back out $u(j)$ that is consistent with $c(j)$ and $\{u'(s'|g)\}_{s'}$ and we use the objective function of the firm, the right hand side of (10) to obtain v_j:

$$u(j) = \left[(1-\beta)\,c(j)^{1-\frac{1}{\psi}} + \beta m(j)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$
$$v(j) = 1 - c(j) - A(\theta(j)) + \kappa\theta(j)\mathbb{E}_g\Lambda(g'|g,x)\, e^{g'+\eta'+\varepsilon'} v(u'(s'|j)\,|g',\Gamma_x(g'|g,x))$$

(g) The guess for $v(u|g,x)$ and $c(u|g,x)$ are updated by interpolating values $\{u_j, v_j\}$ and $\{u_j, c_j\}$. We then iterate until the value function and consumption functions both converge with a tolerance of $1e-7$ under a sup norm.

3. To check the accuracy in computing the optimal contract, we plot a version of Euler equation errors in Figure A1. Fixing u, x, g and the aggregate state next period g', we draw 1000 idiosyncratic shocks $\varepsilon'+\eta'$ such that both agent and firm-side limited commitment constraints are not binding. We then use the maximum absolute log10 ratio of worker's MRS to owners' MRS across these shocks as our measure of Euler Equation Error. We repeat this procedure for different (u, x, g) and g' combinations with values of (u, x) that are not on the grid points where the value function is solved. The Euler equation errors computed this way has the magnitude of -4, which suggests that our approximation is reasonable.

4. We now describe the outer loop where we use optimal policies to simulate the model and update

Γ_x. The details of the simulation procedure are given below:

(a) Let $\phi(t)$ denote the summary measure at time t. In simulations, we approximate the continuous distribution $\phi(t)$ by a finite-state distribution as follows. We choose $u_1^{(t)}, u_2^{(t)}, \cdots, u_{N+1}^{(t)}$, where $u_1^{(t)} = \lambda \overline{u}(g_t, x_t)$ and $u_{N+1}^{(t)} = u^*(g_t, x_t)$. A density ϕ is characterized by a set of grid points $\{\hat{u}[n](t)\}_{n=1}^{N+3}$ and corresponding weights $\{\phi[n](t)\}_{n=1}^{N+3}$ such that

- $\hat{u}[1]$ and $\hat{u}[N+1]$ are the boundaries where the limited commitment constraint binds: $\hat{u}[1] = \lambda\overline{u}(g_t, x_t)$ and $\hat{u}[N+1] = u^*(g_t, x_t)$; $\hat{u}[N+2] = u^*(g_t, x_t)$ is the restarting utility.
- $\{\hat{u}[n]\}_{n=2,3,\cdots N}$ are the interior points: $\hat{u}[j] \in (u_{j-1}, u_j)$, for $j = 2, 3\cdots N$, are chosen appropriately to minimize the approximation error.
- $\phi[1]$ and $\phi[N+1]$ are the income shares of agents with a binding limited commitment constraint at $\hat{u}[1]$ and $\hat{u}[N+1]$, respectively.
- $\{\phi[n]\}_{n=2,3,\cdots N}$ are the income shares of agents in the interior.
- The mass on $\phi[N+2]$ is the share of agents who (re)start at $u^*(\phi, g)$, this include both the newly employed and the new born.
- The mass $\phi[N+3]$, which is the total human capital for the unemployed pool.

(b) Start with an initial distribution of u, denoted $\{\phi_0(u)\}$.

(c) Having solved x_0, use the law of motion of $u'(u, s'|\, g, x)$ to compute ϕ_1. Here we describe a general procedure to solve for $\{\phi[n](t+1);\ \hat{u}[n](t+1);\ x_{t+1}\}_{n=1}^{N+3}$ and B_{t+1} given $\{\phi[n](t);\ \hat{u}[n](t);\ x_t\}_{n=1}^{N+3}$ and B_t. Note that the assumed law of motion gives a natural candidate for x_{t+1}. We denote $x_{t+1} = \Gamma(x_t|\, g_t, g_{t+1})$.

   i. First, we approximate the distribution $s \sim f(\varepsilon + \eta|\, g)$ by a finite dimensional distribution such that $\sum_k^K f_g[j] = 1$ and $\sum_k^K e^{s_k} f_g[j] = 1$, for $g = g_H, g_L$.

   ii. Given $\{\phi[n](t), \hat{u}[n](t)\}_{n=1}^{N+3}$ for period t, conditioning on the realization of aggregate state g_{t+1}, for each $n = 1, 2, \cdots, N+2$, we compute $\{\phi_{t+1}[n,k]\}_{n,k}$. The interpretation is that $\phi_{t+1}[n,k]$ is the total measure of income share that comes from agents with $\hat{u}[n](t)$ and with realization of ε_k, which is given by:

$$\phi_{t+1}[n,k] = (1-\kappa)(1-\theta(\hat{u}[n](t),\ g_t)) f_{g_{t+1}}[k]\, \phi_t[n]\, e^{s_k}, \quad k = 1, 2\ldots, K.$$

   The continuation utility of these agents is $u'(\hat{u}[n](t), g_{t+1}, s_k|\, g_t, x_t)$, a fact that we will use below.

   iii. We now compute $\{\phi_{t+1}[m]\}_{m=1,2,\ldots N+3}$ for the next period. First, compute the measure on the grid points:

$$\phi_{t+1}[1] = \sum_{n=1}^{N+2}\sum_{k=1}^{K} \phi_{t+1}[n,j]\, I_{\{u'(\hat{u}[n](t), g_{t+1}, s_k|g_t, x_t) \le \lambda\overline{u}(g_{t+1}, x_{t+1})\}},$$

$$\phi_{t+1}[2] = \sum_{n=1}^{N+2}\sum_{k=1}^{K} \phi_{t+1}[n,k]\, I_{\left\{u'(\hat{u}[n](t), g_{t+1}, s_k|g_t, x_t) \in \left(u_1^{(t+1)}, u_2^{(t+1)}\right)\right\}},$$

$$\phi_{t+1}[m] = \sum_{n=1}^{N+2}\sum_{k=1}^{K} \phi_{t+1}[n,k]\, I_{\left\{u'(\hat{u}[n](t), g_{t+1}, s_k|g_t, x_t) \in \left[u_{m-1}^{(t+1)}, u_m^{(t+1)}\right)\right\}}, \quad m = 3, \ldots N$$

$$\phi_{t+1}[N+1] = \sum_{n=1}^{N+2}\sum_{k=1}^{K} \phi_{t+1}[n,k]\, I_{\{u'(\hat{u}[n](t), g_{t+1}, s_k|g_t, x_t) \ge u^*(g_{t+1}, x_{t+1})\}},$$

Second, we compute the measure of all restarting agents, which include the newly employed and the new born:

$$\phi_{t+1}[N+2] = \kappa + (1-\kappa)\phi_t[N+3]\chi.$$

Finally, we compute the measure of the unemployed pool:

$$\phi_{t+1}[N+3] = (1-\kappa)\left\{\lambda \sum_{n=1}^{N+2} \theta(\hat{u}[n](t), g_t)\phi_t[n] + [1-\chi]\phi_t[N+3]\right\}$$

iv. The interpretation is again that $\phi[1]$ and $\phi[N+1]$ are the income shares of agents with a binding limited commitment constraint at $\hat{u}[1]$ and $\hat{u}[N+1]$, respectively, and $\{\phi[n]\}_{n=2,3,\cdots N}$ are the income shares of agents in the interior. $\phi[N+2]$ is the income share of agents who enter the employment pool (which include newly employed and the new born), and $\phi_{t+1}[N+3]$ is the share of agents in the unemployed pool.

v. We need to update the vector normalized utilities $\{\hat{u}[n](t+1)\}_{n=1}^{N+2}$. Clearly, we should have $\hat{u}[1](t+1) = \lambda\overline{u}(g_{t+1}, x_{t+1})$, $\hat{u}[N+1](t+1) = u^*(g_{t+1}, x_{t+1})$ and $\hat{u}[N+2](t+1) = u^*(g_{t+1}, x_{t+1})$. For $m = 2, \ldots, N$, we choose $\hat{u}[m](t+1) \in \left[u_{m-1}^{(t+1)}, u_m^{(t+1)}\right)$ such that the resource constraint holds exactly for $u \in \left[u_{m-1}^{(t+1)}, u_m^{(t+1)}\right)$. That is, we pick $\hat{u}[m](t+1)$ to be the solution (denoted $\hat{u}$) to

$$\begin{aligned} &\sum_{n=1}^{N+2}\sum_{k=1}^{K}\phi_{t+1}[n,k]\, c\left(u'\left(\hat{u}[n](t), g_{t+1}, \varepsilon_j \middle| g_t, x_t\right)\middle| g_{t+1}, x_{t+1}\right) I_{\left\{u'\left(\hat{u}[n](t), g_{t+1}, \varepsilon_k \middle| g_t, x_t\right) \in \left[u_{m-1}^{(t+1)}, u_m^{(t+1)}\right)\right\}} \\ = \;& c(\hat{u}| g_{t+1}, x_{t+1})\phi_{t+1}[m]. \end{aligned}$$

vi. Now, we compute B_{t+1}:

$$B_{t+1} = (1-\kappa)[1-\chi(g_{t+1})]B_t + \lambda[1-\kappa]\sum_{m=1}^{N+2}\theta(\hat{u}[m](t)| g_t, x_t)\, b\phi_t[m]$$

(d) Up to now, we have described a procedure to simulate forward the economy. This allows us to compute the market clearing $\left\{x_{t+1}^{MC}\right\}_{t=0}^{\infty}$ as follows:

$$x_{t+1}^{MC} = \sum_{m=1}^{N+2}\phi_{t+1}[m] - \sum_{m=1}^{N+2} c(\hat{u}[m](t+1)| g_{t+1}, x_{t+1})\phi_{t+1}[m] - B_{t+1}. \tag{A3.2}$$

Given the sequence of $\{g_t\}_{t=1}^{T}$, we simulate the economy forward for T periods to obtain $\left\{x_t^{MC}\right\}_{t=0}^{T}$. We divide the sample into four cases: $g_H \to g_H$, $g_H \to g_L$, $g_L \to g_H$, $g_L \to g_L$ and use regression to update the law of motion of x. We go back to step 1 to iterate. Note the under the above procedure, given the sequence of $\{g_t\}_{t=1}^{T}$, the sequence of x_{t+1} that is used for computing decision rules is complete determined by (A3.2). In the simulation, we assume that x_{t+1} follows the perceived law of motion, based on which agent make their decisions. We use the market clearing condition to update the actual law of motion of x and iterate.

(e) We divide the sample into four cases: $g_H \to g_H$, $g_H \to g_L$, $g_L \to g_H$, $g_L \to g_L$ and use regressions (A3.1) to update the law of motion of x. We go back to step 1 to iterate until the unconditional R^2 approaches 99.9%.

A4 Calculation of the replicating portfolio

First, we define $p(u|\phi, g, B)$ to be the present value of a worker's consumption claim normalized by Yh. That is, $p(u|\phi, g, B)$ satisfies

$$\begin{aligned} p(u|\phi, g, B) &= c(u|\phi, g, B) \\ &+\kappa\theta(u|\phi, g, B)\int \Lambda(g'|\phi, g, B)\, e^{g'+\eta'+\varepsilon'} p(u'|\phi', g', B')\,\Omega(ds'|g). \end{aligned}$$

Next, the price-to-dividend ratio of the aggregate stock market, denoted $q(\phi, g, B)$ can be computed as:

$$q(\phi, g, B) = \int v(u|\phi, g, B)\,\phi(du).$$

Let $\mathbf{\Delta}$ be the number of shares and $\mathbf{B}Y$ be the amount of risk-free bond in the replicating portfolio. We denote the next period state variables as

$$\phi_i = \Gamma_\phi(g_i|\phi, g, B), \quad B_i = \Gamma_B(g_i|\phi, g, B), \quad for\ i = H, L.$$

Also, denote the one-period risk-free interest rate as

$$R_f(\phi, g, B) = \frac{1}{\mathbb{E}[\Lambda(g'|\phi, g, B)]}.$$

The replicating portfolio $(\mathbf{\Delta}, \mathbf{B})$ are jointly determined by the two equations, for $i = H, L$,

$$\mathbf{\Delta} q(\phi_i, g_i, B_i)\, x(\phi_i, g_i, B_i)\, e^{g_i} + \mathbf{B} R_f(\phi, g, B) = \int p(u'(u, g_i, \varepsilon'|\phi, g, B)|\phi_i, g_i, B_i)\, f(\varepsilon'|g_i)\, d\varepsilon'.$$

The interpretation is the that the replicating portfolio $(\mathbf{\Delta}, \mathbf{B})$ pays the aggregate component of the agent's consumption in all future periods. Given $(\mathbf{\Delta}, \mathbf{B})$, the share of stocks can be computed as the value of the stock as a fraction of the total value of the replicating portfolio:

$$\Delta_C(u, g, \phi) = \frac{\mathbf{\Delta}[q(\phi, g, B) - 1]\, x(\phi, g, B)}{\mathbf{\Delta}[q(\phi, g, B) - 1]\, x(\phi, g, B) + \mathbf{B}}.$$

A5 More details on wage-pass-through and returns in the cross section

In this section, we provide corroborating evidence for the empirical results in sections 6.2 and 6.3. As a robustness for specification (29), we estimate

$$\begin{aligned} \Delta \log \text{WageBill}_{f,t+1} = \alpha_w + \beta_{w0}\ \text{LaborShare}_{f,t} + \beta_{w1}^{+} \max\{\Delta \log \text{Sales}_{f,t}, 0\} \\ \beta_{w1}^{-} \min\{\Delta \log \text{Sales}_{f,t}, 0\} + \gamma_w^{+} \Delta \max\{\log \text{Sales}_{f,t}, 0\} \times\ \text{LaborShare}_{f,t} \\ + \gamma_w^{-} \min\{\Delta \log \text{Sales}_{f,t}, 0\} \times\ \text{LaborShare}_{f,t} + \lambda_{wt}. \end{aligned} \tag{A5.1}$$

The firm-side limited commitment binds with adverse firm-level shocks. This would imply that γ_w^{-} or the coefficient on the negative part of sales growth should be positive and statistically significant. In table 1, we verify this.

Next we estimate a version of (27) with total assets and book leverage as further controls. In table 2, we verify that the coefficient on labor leverage remains positive and statistically significant.

References

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

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

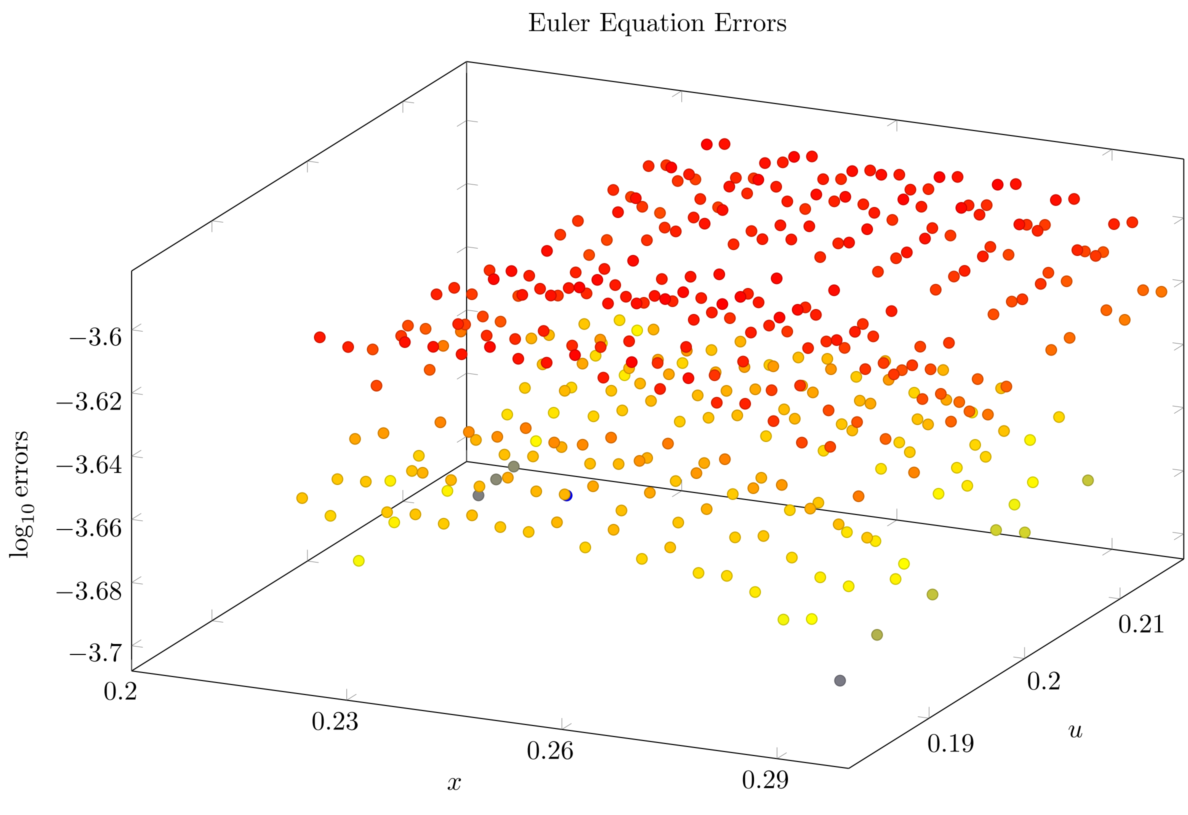

Figure A1: Euler equation errors for $g = g_L$ and $g' = g_L$.

Table 1: FIRM-LEVEL WAGE PASS-THROUGHS AND LABOR SHARES

Coefficients	Using LS	Using ELS
LogSales_plus	0.69	0.55
	(0.05)	(0.04)
LogSales_minus	0.13	0.12
	(0.06)	(0.04)
Labor share	-0.02	-0.12
	(0.01)	(0.01)
Labor share × LogSales_plus	0.00	0.1
	(0.12)	(0.06)
Labor share × LogSales_minus	0.76	0.55
	(0.12)	(0.08)
Time Fixed Effects	Yes	Yes

Notes: The sample consist of firm-year observations from COMPUSTAT/CRSP merged files for the years 1959-2016. In the column labeled "Using LS" we use labor share computed using (28), 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.

Table 2: FIRM-LEVEL RETURNS AND LABOR SHARES

Coefficients	Using LS	Using ELS
Labor share	1.24	0.85
	(0.43)	(0.20)
Leverage	0.63	1.17
	(0.91)	(0.35)
log Assets	-1.06	-2.22
	(0.43)	(0.21)
Time Fixed Effects	Yes	Yes

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. *Log Assets* is the logarithm of book value of assets and *Leverage* is defined as the ratio of long-term debt plus debt in current liabilities divided by total assets.