MIT Libraries | DSpace@MIT

MIT Open Access Articles

Left Behind: Creative Destruction, Inequality, and the Stock Market

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Kogan, Leonid et al. "Left Behind: Creative Destruction, Inequality, and the Stock Market." Journal of Political Economy 128, 3 (March 2020): 855-906 © 2020 by The University of Chicago

Published Version: https://doi.org/10.1086/704619

Publisher: University of Chicago Press

Permanent Link: https://hdl.handle.net/1721.1/130433

Version: Final published version

Terms of use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Massachusetts Institute of Technology

Left Behind: Creative Destruction, Inequality, and the Stock Market

Leonid Kogan

Massachusetts Institute of Technology and National Bureau of Economic Research

Dimitris Papanikolaou

Kellogg School of Management and National Bureau of Economic Research

Noah Stoffman

Kelley School of Business

We develop a general equilibrium model of asset prices in which benefits of technological innovation are distributed asymmetrically. Financial market participants do not capture all economic gains from innovation even when they own shares in innovating firms. Such gains accrue partly to the innovators, who cannot sell claims on proceeds from their future ideas. We show how the resulting inequality among agents can give rise to a high risk premium on the aggregate stock market, return comovement and average return differences among firms, and the failure of traditional representative agent asset pricing models to account for cross-sectional differences in risk premia.

I. Introduction

Technological innovation is arguably the main driver of economic growth in the long run. However, the economic value generated by new ideas is

Earlier versions of this paper circulated under the titles "Technological Innovation: Winners and Losers" and "Winners and Losers: Creative Destruction and the Stock Market." We thank Andrew Abel, Hengjie Ai, Federico Belo, Jaroslav Borovička, John Campbell, Ian Dew-Becker, Bernard Dumas, Andrea Eisfeldt, Carola Frydman, Joao Gomes, Lars Hansen, Boyan Jovanovic, Erik Loualiche, Deborah Lucas, Gregor Matvos, Stavros Panageas, Jonathan

Electronically published January 23, 2020
[*Journal of Political Economy*, 2020, vol. 128, no. 3]
© 2020 by The University of Chicago. All rights reserved. 0022-3808/2020/12803-0007$10.00

usually not shared equally. The popular press is rife with rags-to-riches stories of new entrepreneurs whose net worth rose substantially as a result of their innovative ideas during technological booms, such as those experienced in the 1990s and the 2000s. In addition to the large wealth gains for successful innovators, technological progress can also create losses through creative destruction, as new technologies render old capital and processes obsolete.

We show that the asymmetric sharing of gains and losses from technological innovation can give rise to well-known, prominent empirical patterns in asset price behavior, including a high risk premium on the aggregate stock market, return comovement, and average return differences among growth and value firms. We build a tractable general equilibrium model in which the benefits of technological progress are distributed unevenly across investors and firms. Our model allows for two forms of technological progress. Some advances take the form of improvements in labor productivity and are complementary to existing investments, while others are embodied in new vintages of capital. Throughout the paper, we refer to the first type of technological progress as disembodied and the second type as embodied.[1] The latter type of technological progress leads to more creative destruction, since old and new capital vintages are substitutes.

Parker, Monika Piazzesi, Nick Roussanov, Martin Schneider, Amit Seru, and Costis Skiadas as well as seminar participants at American Economic Association, Baruch College, the BI Production–based asset pricing workshop, Boston University, Catolica, Chicago Booth, Chicago Institute for Theoretical Economics, Columbia Graduate Business School, Copenhagen Business School, Chinese University of Hong Kong, Emory, Erasmus University, Georgetown University, Georgia Institute of Technology, École des Hautes Études Commerciales de Paris, Hong Kong University of Science and Technology, Institut Européen d'Administration des Affaires, Jackson Hole Finance Conference, London Business School, London School of Economics, McGill University, Monash University, Massachusetts Institute of Technology Sloan, National Bureau of Economic Research Asset Pricing meeting, New York Federal Reserve Bank, Northwestern University, Rice University, Society of Economic Dynamics, Stockholm School of Economics, Tinbergen Institute, University of California, Berkeley, University of California, Los Angeles, University of Hong Kong, University of Illinois at Chicago, Universidad de San Andres, Universidad Torcuato Di Tella, University of Connecticut, University of Geneva, Indiana University, University of Lausanne, University of Lugano, University of New South Wales, University of Rochester, University of Sydney, University of Technology Sydney, University of Washington, University of Southern California, Washington University, Western Finance Association, and Wharton for valuable discussions and comments. Dimitris Papanikolaou thanks Amazon (Amazon Web Services in Education Grant award), the Zell Center for Risk, and the Jerome Kenney Fund for research support. Data are provided as supplementary material online

[1] Berndt (1990, 71) gives the following definitions for these two types of technology shocks: "Embodied technical progress refers to engineering design and performance advances that can only be embodied in new plant or equipment; older equipment cannot be made to function as economically as the new, unless a costly remodelling or retrofitting of equipment occurs"; and "By contrast, disembodied technical progress refers to advances in knowledge that make more effective use of all inputs, including capital of each surviving vintage (not just the most recent vintage). In its pure form, disembodied technical progress proceeds independently of the vintage structure of the capital stock. The most common example of disembodied technical progress is perhaps the notion of learning curves, in which it has been found that for a wide variety of production processes and products, as cumulative experience and production increase, learning occurs which results in ever decreasing unit costs."

A prominent feature of our model is that the market for new ideas is incomplete. Specifically, shareholders cannot appropriate all the economic rents generated by new technologies, even when they own equity in the firms that develop those technologies. Our motivation for this market incompleteness is that ideas are a scarce resource, and the generation of ideas relies heavily on human capital. As a result, innovators are able to capture a fraction of the economic rents that their ideas generate. The key friction is that potential innovators cannot sell claims to these future rents. This market incompleteness implies that technological progress has an asymmetric impact on household wealth. Most of the financial benefits from innovation accrue to a small fraction of the population, while the rest bear the cost of creative destruction. This reallocative effect of technological progress is particularly strong when innovations are embodied in new capital goods. By exposing households to idiosyncratic randomness in innovation outcomes, improvements in technology can thus reduce households' indirect utility. This displacive effect on indirect utility is amplified when households also care about their consumption relative to the economy-wide average, since households dislike being left behind.

Displacement risk contributes to the equity risk premium and also leads to cross-sectional differences in asset returns. Owning shares in growth firms helps offset potential utility losses brought on by technological improvements. In our model, firms differ in their ability to acquire projects that implement new technologies. This difference in future growth opportunities implies that technological progress has a heterogeneous impact on the cross section of asset returns. Firms with few existing projects—but many potential new investment opportunities—benefit from technological advances. By contrast, profits of firms that are heavily invested in old technologies and have few growth opportunities decline because of increased competitive pressure. In equilibrium, investors hold growth firms, despite their lower expected returns, as a hedge against the potential wealth reallocation that may result from future technological innovation. Aggregate consumption does not accurately reflect all the risks that households face as a result of technological progress, implying the failure of traditional representative agent asset pricing models to account for cross-sectional differences in risk premia.

We estimate the parameters of the model using indirect inference. The baseline model performs well at replicating the joint properties of aggregate consumption, investment, and asset returns that we target. In particular, the model generates an aggregate consumption process with moderate low-frequency fluctuations, volatile equity returns, a high equity premium, a low and stable risk-free rate, and the observed differences in average returns between value and growth stocks (the value premium). Jointly replicating these patterns has proven challenging in existing representative agent general equilibrium models. These results depend on three key features of our model: technology advances that are embodied in new capital goods,

an incomplete market for ideas, and preferences over relative consumption. Restricted versions of the model that eliminate any one of these three features have difficulty replicating the empirical properties of asset returns, especially the cross-sectional differences in stock returns between value and growth firms. In sum, these three features are responsible for generating sufficient displacement risk in the model and ensuring that households care sufficiently about displacement to endure that the embodied shock carries a negative risk price.

In addition to the features of the data that we target, we evaluate the model's other implications for asset returns. The model generates realistic predictions about income and wealth inequality, especially at the top of the distribution. Importantly, the model replicates the observed patterns of comovement between value and growth stocks (the value factor) and the failure of the capital asset pricing model (CAPM) and the consumption CAPM to explain the value premium.[2] In addition, the model generates an upward-sloping real yield curve and a downward-sloping term structure of risk premia on corporate payout strips.

In the last part of the paper, we provide additional evidence that directly relates to the main mechanism in the model. Most of the model's predictions rely on the fluctuations in the share of economic value that is generated by new innovative ideas, or blueprints—a quantity that is challenging to measure empirically. We construct an empirical estimate of this share using data on patents and stock returns collected by Kogan et al. (2017). Following Kogan et al. (2017), we infer the economic value of patents from the firms' stock market reaction following a successful patent application.[3]

[2] The value puzzle consists of two robust empirical patterns. First, firms with higher than average valuations—growth firms—experience lower than average future returns. These differences in average returns are economically large and comparable in magnitude to the equity premium. This finding has proven to be puzzling because growth firms are typically considered to be riskier and therefore should command higher average returns. Existing asset pricing models, such as the CAPM and the consumption CAPM, largely fail to price the cross section of value and growth firms. Second, stock returns of firms with similar valuation ratios exhibit comovement, even across industries. These common movements are typically unrelated to the firms' exposures to fluctuations in the overall market value. See Fama and French (1992, 1993) for more details. Patterns similar to the returns of high market-to-book firms have been documented for firms with high past investment (Titman, Wei, and Xie 2004) or price-earnings ratios (Rosenberg, Reid, and Lanstein 1985). The strong patterns of return comovement among firms with similar characteristics have motivated the use of empirical factor models (Fama and French 1993). However, the economic origins of these empirical return factors are yet to be fully understood. Our work can thus be viewed as providing a microfoundation for including the value factor in reduced-form asset pricing models.

[3] Relative to other measures of innovation, such as patent citations, the stock market reaction to patent issues has the unique advantage of allowing us to infer the economic—as opposed to the scientific—value of the underlying innovations. Focusing on the days around the patent is issued allows us to infer the economic value over a narrow time window. However,

The correlations between the estimated value of new blueprints and aggregate quantities and prices are largely consistent with the model. Increases in the value of new blueprints are associated with higher aggregate investment, lower market returns, and higher returns for growth firms relative to value firms. Further, increases in the relative value of new blueprints are associated with increases in consumption, income, and wealth inequality. Consistent with the model, high-Q firms are more likely to innovate in the future than low-Q firms. Last, we find that rapid technological progress within an industry is associated with lower future profitability for low Tobin's Q (value) firms relative to high-Q (growth) firms. We replicate these results in simulated data from the model; the empirical estimates are in most cases close to those implied by the model. We interpret these findings as providing support for the model's main mechanism.

In sum, the main contribution of our work is to develop a general equilibrium model that introduces a new mechanism: it relates increases in inequality following technological innovations to the pricing of shocks to technology in financial markets.[4] Our work thus adds to the growing literature studying asset prices in general equilibrium models.[5] Conceptually closest to our paper is that of Garleanu, Kogan, and Panageas (2012), who study the value premium puzzle in an overlapping generations economy

---

the stock market reaction may underestimate the value of the patent when some of the information about the innovation is already priced in by the market. Further, our analysis misses the patents issued to private firms as well as those inventions that are not patented. Our analysis should thus be viewed as a joint test of the model and the assumption that movements in the economic value of patented innovations are representative of the fluctuations in the overall value of new inventions in the economy.

[4] Kogan and Papanikolaou (2013, 2014) feature a similar structural model of the firm in a partial equilibrium setup. Partial equilibrium models are useful in connecting factor risk exposures to firm characteristics. However, they take the stochastic discount factor (SDF) as given. Reduced-form specifications of the SDF can arise in economies in which all cross-sectional variation in expected returns is due to sentiment (see, e.g., Nagel, Kozak, and Santosh 2014). A general equilibrium model is necessary to connect the SDF to the real side of the economy.

[5] Related to this paper, Papanikolaou (2011) and Garleanu, Panageas, and Yu (2012) feature representative agent models that also include capital-embodied technology shocks. Papanikolaou (2011) focuses on the pricing of embodied shocks in an environment with a representative firm and complete markets. In his model, capital-embodied shocks carry a negative risk premium due to agents' aversion to short-run consumption fluctuations. The mechanism of Papanikolaou (2011) requires that positive capital-embodied shocks are associated with reductions in aggregate consumption on impact. The empirical evidence for this is mixed. By contrast, we propose a different mechanism that leads to a negative risk price for embodied shocks. Garleanu, Panageas, and Yu (2012) focus on understanding the joint time series properties of consumption and excess asset returns; our main focus is on the model's cross-sectional implications. Closely related work also features heterogeneous firms. Gomes, Kogan, and Zhang (2003) study the cross section of risk premia in a model where technology shocks are complementary to all capital. Pastor and Veronesi (2009) study the pricing of technology risk in a model with a life cycle of endogenous technology adoption. Ai, Croce, and Li (2013) analyze the value premium in a model where some technology shocks affect the productivity of only old capital. However, all of these studies consider representative agent models.

where technological improvements lead to intergenerational displacement risk.[6]

Our model contains features that connect it to several other strands of the literature. A key part of our model mechanism is that technological progress endogenously increases households' uninsurable consumption risk. The fact that time-varying cross-sectional dispersion of consumption can increase the volatility of the SDF is well known (Constantinides and Duffie 1996). However, whereas the existing literature uses reduced-form specifications of idiosyncratic labor income risk, in our setting the time variation in households' uninsurable risk arises as an equilibrium outcome. The resulting effect of idiosyncratic risk on asset prices is further amplified by households' preferences over relative consumption. Our work thus builds on the extensive literature that emphasizes the role of consumption externalities and relative wealth concerns for asset prices, investment, and consumption dynamics (Duesenberry 1949; Abel 1990; Roussanov 2010). Closest to our work is that of Roussanov (2010), who argues that households may invest in risky, zero-mean gambles whose payoffs are uncorrelated with the aggregated state when they have preferences over their rank in the consumption distribution. In our setting, preferences over relative consumption induce agents to accept low-risk premia (or equivalently high valuations) to hold assets that increase in value when technology prospects improve—that is, growth firms. Last, the idea that a significant fraction of the rents from innovation accrue to human capital is related to Atkeson and Kehoe (2005), Lustig, Syverson, and Van Nieuwerburgh (2011) and Eisfeldt and Papanikolaou (2013). In contrast to these authors, we explore how fluctuations in the value of these rents affect the equilibrium pricing of technology shocks.

[6] Our paper shares some of the elements in their model, namely, incomplete markets. Specifically, Garleanu, Kogan, and Panageas (2012) show that consumption inequality is an intergenerational phenomenon; there is no inequality in consumption within household cohorts. This stylized assumption makes the model analytically tractable but makes it difficult to reconcile the model with the patterns of inequality in the data. For instance, income inequality within cohorts is typically much larger than inequality between cohorts (O'Rand and Henretta 2000). Also, Song et al. (2018) document that most of the rise in income inequality over the 1978–2012 period is within cohorts. In section C.3 in the appendix (available online), we perform a variance decomposition exercise and show that most of the inequality in consumption, income, or wealth exists within cohorts. Allowing for intracohort heterogeneity is a nontrivial analytical challenge because models of individual heterogeneity with incomplete markets and aggregate shocks typically do not aggregate to a finite dimensional state space. Another contribution of our paper is thus to provide a tractable model with individual heterogeneity and aggregate shocks that has realistic implications for inequality. Unlike Garleanu, Kogan, and Panageas (2012), our model also incorporates firm investment and generates a realistic (and stationary) cross-sectional distribution of firms. In sum, the fact that the model has realistic predictions about inequality and firm heterogeneity allows for a more rigorous evaluation of the quantitative importance of the mechanism.

II. The Model

We consider a dynamic continuous-time economy, with time indexed by t. We first introduce the productive sector of the economy: firms and the projects they own. We next introduce households and describe the nature of market incompleteness in our setup.

A. *Firms and Technology*

The basic production unit in our economy is called a project. Projects are owned and managed by firms. Each firm hires labor services to operate the projects it owns. The total output of all projects can be used to produce either consumption or investment. New production units are created using investment goods and project blueprints (ideas). Households supply labor services and blueprints to firms and derive utility from consumption. Figure 1 summarizes the structure of our model.

1. Active Projects

Each firm f owns a constantly evolving portfolio of projects, which we denote by $\mathcal{J}_{ft}$. We assume that there is a continuum of infinitely lived firms in the economy, which we index by $f \in [0, 1]$. Projects are differentiated from each other by three characteristics: (1) their operating scale, determined by

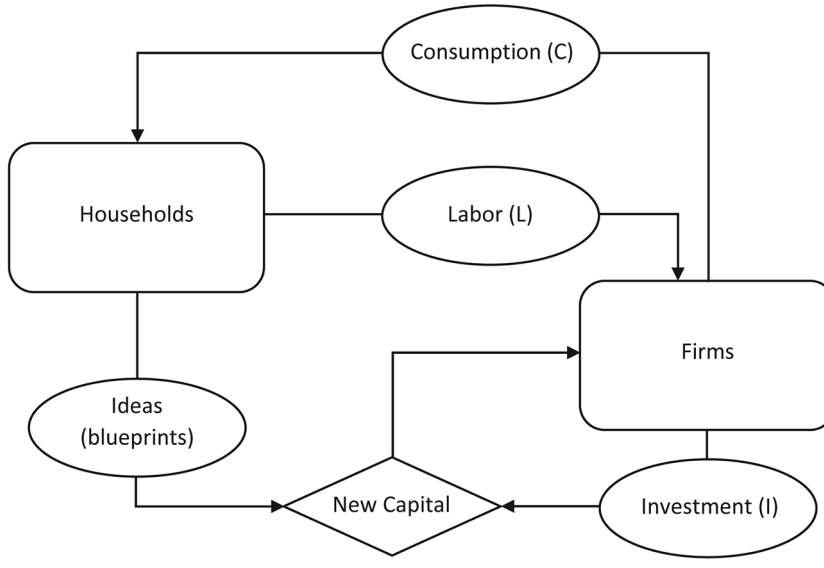


FIG. 1.—Production and investment flow chart.

the amount of capital goods associated with the project, k; (2) the systematic component of project productivity, ξ; and (3) the idiosyncratic (or project-specific) component of productivity, u. Project j, created at time $\tau(j)$, produces a flow of output equal to

$$Y_{j,t} = \left(u_{j,t}\, e^{\xi_{\tau(j)}} K_{j,t}\right)^{\phi} \left(e^{x_t} L_{j,t}\right)^{1-\phi}, \tag{1}$$

where $L_{j,t}$ is labor allocated to project j. In contrast to the scale decision, the choice of labor allocated to project $L_{j,t}$ can be freely adjusted every period. Firms purchase labor services at the equilibrium wage w_t. We denote by

$$\Pi_{j,t} = \sup_{L_{j,t}} \left[\left(u_{j,t}\, e^{\xi_{\tau(j)}} K_{j,t}\right)^{\phi} \left(e^{x_t} L_{j,t}\right)^{1-\phi} - w_t L_{j,t} \right] \tag{2}$$

the profit flow of project j under the optimal hiring policy.

We emphasize one important dimension of heterogeneity among technological innovations by modeling technological progress using two independent processes, ξ_t and x_t. First, the shock ξ reflects technological progress embodied in new projects. It follows an arithmetic random walk

$$d\xi_t = \mu_\xi dt + \sigma_\xi dB_{\xi,t}, \tag{3}$$

where B_ξ is a standard Brownian motion. The term ξ_s denotes the level of frontier technology at time s. Growth in ξ affects only the output of new projects created using the latest frontier of technology. In this respect, our model follows the standard vintage capital model of Solow (1960).

Second, the labor-augmenting productivity process x_t follows an arithmetic random walk

$$dx_t = \mu_x dt + \sigma_x dB_{x,t}. \tag{4}$$

Here, B_x is a standard Brownian motion independent of all other productivity shocks. In particular, the productivity process x is independent from the embodied productivity process ξ. Labor in our model is complementary to capital. Thus, in contrast to the embodied shock ξ, the technology shock x affects the output of all vintages of existing capital.

The level of project-specific productivity u_j is a stationary mean-reverting process that evolves according to

$$du_{j,t} = \kappa_u(1 - u_{j,t})\, dt + \sigma_u\, u_{j,t} dB^u_{j,t}, \tag{5}$$

where B^u_j are standard Brownian motions independent of B_ξ. We assume that $dB^u_{j,t} \cdot dB^u_{j',t} = dt$ if projects j and j' belong in the same firm f and zero otherwise. As long as $2\kappa_u \geq \sigma_u^2$, the ergodic distribution of u has finite first two moments (for details, see lemma 1 in the appendix). All new projects implemented at time s start at the long-run average level of idiosyncratic productivity, that is, $u_{j,\tau(j)} = 1$. Thus, all projects created at a point

in time are ex ante identical in terms of productivity but differ ex post because of the project-specific shocks.

The firm chooses the initial operating scale k of a new project irreversibly at the time of its creation. Firms cannot liquidate existing projects and recover their investment costs. Over time, the scale of the project diminishes according to

$$dK_{j,t} = -\delta K_{j,t}dt, \tag{6}$$

where δ is the economy-wide depreciation rate. At this stage, it is also helpful to define the aggregate stock of installed capital, adjusted for quality:

$$K_t = \int_0^1 \left(\sum_{j \in \mathcal{J}_{f,t}} e^{\xi_{\tau(j)}} K_{j,t} \right) df. \tag{7}$$

The aggregate capital stock K also depreciates at rate δ.

2. Creation of New Projects

Creating a new project requires a blueprint and new investment goods. Firms are heterogeneous in their ability to acquire new blueprints. Inventors initially own the blueprints for creation of new projects. We assume that inventors lack the ability to implement these ideas on their own and instead sell the blueprints for new projects to firms (we outline the details of the process for blueprint sales below).

Firms acquire projects by randomly meeting inventors who supply blueprints. The likelihood of acquiring a new project is exogenous to each firm, driven by a doubly stochastic Poisson process $N_{f,t}$ with increments independent across firms. The arrival rate of new projects equals $\lambda_{f,t}$. This arrival rate is time varying and follows a two-state continuous-time Markov chain with high and low growth states $\{\lambda_H, \lambda_L\}$, $\lambda_H > \lambda_L$. The transition rate matrix is given by

$$\begin{pmatrix} -\mu_L & \mu_L \\ \mu_H & -\mu_H \end{pmatrix}. \tag{8}$$

We denote the unconditional average of $\lambda_{f,t}$ by λ. The stochastic nature of λ_{ft} has no effect on the aggregate quantities in the model. Parameters of this process control the cross-sectional differences in investment, valuation ratios, and systematic risk among firms.

To implement a new blueprint as a project j at time t, a firm purchases new capital goods in quantity $I_{j,t}$. Investment in new projects is subject to decreasing returns to scale,

$$K_{j,t} = I_{j,t}^{\alpha}. \tag{9}$$

The parameter $\alpha \in (0, 1)$ parameterizes the investment cost function and implies that costs are convex at the project level.

The value of new ideas (blueprints) plays a key role in our analysis. We denote by

$$\nu_t \equiv \sup_{K_{j,t}} \left\{ \mathrm{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} \Pi_{j,s} \, ds \right] - K_{j,t}^{1/\alpha} \right\} \tag{10}$$

the net value of a new project implemented at time t under the optimal investment policy, where Λ_t is the equilibrium SDF defined in section II.D. Since all projects created at time t are identical ex ante, ν is independent of j. Equation (10) is also equal to the value of a new blueprint associated at time t. When we take the model to the data, we will proxy for the value of new projects (eq. [10]) using the estimates of the value of new patents based on the methodology of Kogan et al. (2017).

3. Aggregate Output

The total output in the economy is equal to the aggregate output of all active projects,

$$Y_t = \int_0^1 \left(\sum_{j \in \mathcal{J}_{f,t}} Y_{j,t} \right) df. \tag{11}$$

The aggregate output of the economy can be allocated to either investment I_t or consumption C_t,

$$Y_t = I_t + C_t. \tag{12}$$

The amount of new investment goods I_t produced is used as an input in the implementation of new projects, as given by the investment cost function defined in equation (9).

B. Households

There is a continuum of households, with the total measure of households normalized to one. Households die independently of each other according to the first arrival of a Poisson process with arrival rate δ^h. New households are born at the same rate, so the total measure of households remains constant. All households are endowed with the unit flow rate of labor services, which they supply inelastically to the firms producing the final good.

Households have access to financial markets and optimize their lifetime utility of consumption. Households are not subject to liquidity constraints; hence, they sell their future labor income streams and invest the

proceeds in financial claims. We denote consumption of an individual household i by $C_{i,t}$.

All shareholders have the same preferences, given by

$$J_t = \lim_{\tau \to \infty} \mathrm{E}_t \left[\int_t^\tau \Upsilon(C_s, J_s; \bar{C}_s) ds \right], \tag{13}$$

where $\Upsilon(C, J; \bar{C})$ is the utility aggregator:

$$\Upsilon(C, J; \bar{C}) = \frac{\rho}{1 - \theta^{-1}} \left\{ \frac{\left[C^{1-h} (C/\bar{C})^h \right]^{1-\theta^{-1}}}{\left[(1-\gamma) J \right]^{(\gamma - \theta^{-1})/(1-\gamma)}} - (1-\gamma) J \right\}. \tag{14}$$

Households' preferences fall into the class of stochastic differential utility proposed by Duffie and Epstein (1992), which is a continuous-time analog of the preferences proposed by Epstein and Zin (1989). Relative to Duffie and Epstein (1992), our preference specification also incorporates a relative consumption concern (keeping up with the Joneses; see, e.g., Abel 1990). That is, households also derive utility from their consumption relative to aggregate consumption,

$$\bar{C} = \int_0^1 C_{i,t} di. \tag{15}$$

The parameter h captures the strength of the relative consumption effect; γ is the coefficient of relative risk aversion; θ is the elasticity of intertemporal substitution (EIS); and ρ is the effective time-preference parameter, which includes the adjustment for the likelihood of death.

C. *Household Innovation*

The key feature of our model is imperfect risk sharing among investors. Households are endowed with ideas, or blueprints, for new projects. Inventors do not implement these project on their own. Instead, they sell the ideas to firms. Inventors and firms bargain over the surplus created by new projects; the inventor captures a share η of the net present value of a new project.

Each household receives blueprints for new projects according to an idiosyncratic Poisson process with arrival rate μ_I. In the aggregate, households generate blueprints at a rate equal to the total measure of projects acquired by firms, λ. Not all innovating households receive the same measure of new blueprints. Each household i receives a measure of projects in proportion to its wealth $W_{i,t}$, that is, equal to $\lambda W_{i,t} (\mu_I \int_0^1 W_{n,t}\, dn)^{-1}$. This is a technical assumption that is important for tractability of the model, as we discuss below. Thus, conditional on innovating, wealthier households receive a larger measure of blueprints.

Importantly, households cannot trade in securities contingent on future successful individual innovation. That is, they cannot sell claims against their proceeds from future innovations. This restriction on risk sharing plays a key role in our setting. In equilibrium, wealth creation from innovation leads to changes in the cross-sectional distribution of wealth and consumption and therefore affects households' financial decisions.

D. Financial Markets

We assume that agents can trade a complete set of state-contingent claims contingent on the paths of the aggregate and idiosyncratic productivity processes as well as paths of project arrival rates and project arrival events at the firm level. We denote the equilibrium SDF by Λ_t. In addition, we follow Blanchard (1985) and assume that investors have access to competitive annuity markets that allow them to hedge their mortality risk. This assumption implies that, conditional on surviving during the interval $[t, t + dt]$, investor i collects additional income proportional to her wealth, $\delta^h W_{i,t} dt$.

E. Discussion of the Model's Assumptions

Most existing production economy general equilibrium models of asset returns build on the neoclassical growth framework. We depart from this literature in three significant ways.

1. Technological Progress Is Embodied in New Capital Vintages

Most existing general equilibrium models that study asset prices assume that technological progress is complementary to the entire existing stock of capital, as is the case for the x shock in our model. However, many technological advances are embodied in new capital goods and thus benefit only firms that invest in the new capital vintages. Several empirical studies show substantial vintage effects in plant productivity. For instance, Jensen, McGuckin, and Stiroh (2001) find that the 1992 cohort of new plants was 50% more productive than the 1967 cohort in their respective entry years, controlling for industry-wide factors and input differences. Further, an extensive literature documents a significant impact of embodied technological progress on economic growth and fluctuations (see, e.g., Greenwood, Hercowitz, and Krusell 1997; Fisher 2006). Since technological progress can take many forms, we distinguish between embodied and disembodied technological progress to obtain a broader understanding of how technological change affects asset returns.

In our model, capital-embodied technical change introduces a wedge between the value of installed capital and the value of future growth opportunities. This allows our model to deliver a realistic degree of heterogeneity in investment, valuations, and stock returns among firms. Furthermore, because of their heterogeneous impact on the different sources of wealth, embodied shocks combined with incomplete markets help generate realistic levels of consumption inequality within the model.

2. Incomplete Markets for Innovation

Incomplete markets play a key role in our analysis. In the model, inventors generate new ideas and sell them to firms. Importantly, the economic value that is generated by new ideas cannot be fully pledged to outside investors. The goal of this assumption is to allow for a heterogeneous impact of technology shocks on households, which implies that the technology risk exposures of individual investors are not fully captured by aggregate consumption dynamics.

We should emphasize that our interpretation of inventors is quite broad. For example, the term "inventors" can include highly skilled personnel, who generate new inventions or business ideas; entrepreneurs and startup employees, who can extract a large share of the surplus created by new ideas; angel investors and venture capitalists, who help bring these ideas to market; and corporate executives, who decide how to optimally finance and implement these new investment opportunities. In sum, inventors in our model encapsulate all parties that share the rents from new investment opportunities besides the owners of the firm's publicly traded securities. Further, the exact process by which inventors and shareholders share the rents from new technologies can take many forms. One possibility is that inventors work for existing firms, generate ideas, and receive compensation commensurate with the economic value of their ideas. Since their talent is in scarce supply, these skilled workers may be able to capture a significant fraction of the economic value of their ideas. Another possibility is that inventors implement the ideas themselves, creating startups that are partly funded by outside investors. Innovators can then sell equity in these startups to investors and thus capture a substantial share of the economic value of their innovations.[7]

The assumption that households cannot share rents from innovation ex ante can be motivated on theoretical grounds. New ideas are the product of human capital, which is inalienable. Hart and Moore (1994) show

[7] Possibly, the threat of implementing the ideas themselves may be what ensures that existing employees can appropriate some fraction of the benefits. According to Bhide (1999), 71% of the founders of firms in the Inc. 500 list of fast-growing technology firms report that they replicated or modified ideas encountered through previous employment.

that the inalienability of human capital limits the amount of external finance that can be raised by new ventures. Bolton, Wang, and Yang (2015) characterize a dynamic optimal contract between a risk-averse entrepreneur with risky inalienable human capital and firm investors. The optimal contract involves a trade-off between risk sharing and incentives and leaves the entrepreneur with a significant fraction of the upside.

3. Preferences over Relative Consumption

Quantitative asset pricing models often assume relative consumption preferences, which help increase household's aversion to consumption growth volatility and thus raise the equilibrium compensation for bearing consumption risk (Abel 1990; Constantinides 1990; Campbell and Cochrane 1999). In our model, relative consumption preferences make households averse to consumption inequality, thus magnifying the effect of limited sharing of innovation risk.

The assumption of relative consumption preferences is theoretically appealing and has direct empirical support. Rayo and Becker (2007) propose a theory in which peer comparisons are an integral part of the happiness function as a result of an evolutionary process. DeMarzo, Kaniel, and Kremer (2008) show that competition over scarce resources can make agents' utilities dependent on the wealth of their cohort. In particular, if households have preferences over a consumption basket $\hat{C} = C^{1-\kappa}H^{\kappa}$, where C is nondurable consumption and H is a consumption good that is in inelastic supply—for instance, land or local services, such as education—then the household makes intertemporal choices as if it has preferences over the composite good $\hat{C}' = C^{1-\kappa}w^{\kappa}$, where w denotes the household's wealth share.

In a series of seminal papers, Easterlin notes that income and self-reported happiness are positively correlated across individuals within a country but that average happiness within countries does not seem to rise over time as countries become richer (see, e.g., Easterlin 1974). Easterlin interprets these findings as evidence that relative—rather than absolute—income matters for well-being. Consistent with this view, several empirical studies document that, controlling for household income, income of a peer group is negatively related to self-reported measures of happiness and satisfaction (Clark and Oswald 1996; Luttmer 2005; Card et al. 2012). These relative income concerns are substantial. For example, the point estimates of Luttmer (2005) imply that the income of households in the same metropolitan area is more important for happiness than the households' own level of income. Frydman (2015) finds direct evidence for utility preferences over relative wealth in an experimental setting using neural data collected through functional magnetic resonance imaging.

In section II.E, we describe the sensitivity of our quantitative results to assumptions 1–3. In addition to these three main assumptions, our model deviates in some other respects from the neoclassical framework. These deviations make the model tractable but do not drive our main results.

First, we assume that projects arrive independently of the firms' own past decisions and that firms incur convex investment costs at the project level. Together, these assumptions ensure that the optimal investment decision can be formulated as a static problem, thus implying that the cross-sectional distribution of firm size does not affect equilibrium aggregate quantities and prices. Second, the assumption that innovating households receive a measure of projects that is proportional to their existing wealth, together with homotheticity of preferences, implies that all households are solving the same consumption-portfolio choice problem. In equilibrium, households do not trade with each other—similar to Constantinides and Duffie (1996)—and their optimal consumption and portfolio choices scale in proportion to their wealth. The cross-sectional distribution of household wealth then does not affect equilibrium prices.[8] Third, households in our model have finite lives. This assumption ensures the existence of a stationary distribution of wealth among households. Fourth, our assumption that project productivity shocks are perfectly correlated at the firm level ensures that the firm state vector is low dimensional. Last, there is no cross-sectional heterogeneity among the quality of different blueprints. We could easily allow for an idiosyncratic part to ξ—perhaps allowing for substantial skewness in this component—to capture the notion that the distribution of profitability of new ideas can be highly asymmetric. Our conjecture is that such an extension would strengthen our main results by raising the level of idiosyncratic risk of individual households' consumption processes.

F. Competitive Equilibrium

Here we describe the competitive equilibrium of our model. Our equilibrium definition is standard and is summarized below.

DEFINITION 1 (Competitive equilibrium). The competitive equilibrium is a sequence of quantities $\{C_t, I_t, Y_t, K_t\}$, prices $\{\Lambda_t, \mathrm{w}_t\}$, household consumption decisions $\{C_{i,t}\}$, and firm investment and hiring decisions $\{I_{j,t}, L_{j,t}\}$ such that, given the sequence of stochastic shocks $\{x_t, \xi_t, u_{j,t}, N_{f,t}\}$, $j \in \cup_{f\in[0,1]}\mathcal{J}_{f,t}$, $f \in [0, 1]$: (1) households choose consumption and savings plans to maximize

[8] The assumption that the magnitude of innovation is proportional to households' wealth levels likely weakens our main results compared with the case where all households received the same measure of blueprints upon innovating. In the latter case, wealthier households would benefit less from innovation, raising their exposure to innovation shocks relative to our current specification.

their utility (eq. [13]); (2) household budget constraints are satisfied; (3) firms maximize profits; (4) the labor market clears, $\int_0^1 (\Sigma_{j \in \mathcal{J}_{f,t}} L_{j,t}) df = 1$; (5) the demand for new investment equals supply, $\int_0^1 I_{n,t} dn = I_t$; (6) the market for consumption clears, $\int_0^1 C_{n,t} dn = C_t$; and (7) the aggregate resource constraint (11) is satisfied.

We solve for equilibrium prices and quantities numerically. Because of market incompleteness, standard aggregation results do not apply. Specifically, there are two dimensions of heterogeneity in the model: on the supply side among firms and on the demand side among households. Both of these sources of heterogeneity can potentially make the state space of the model infinite dimensional. However, the first two assumptions discussed in section II.E enable a relatively simple characterization of equilibrium; we can solve for aggregate-level quantities and prices in definition 1 as functions of a low-dimensional Markov aggregate state vector.

Specifically, the first moment of the cross-sectional distribution of installed capital K summarizes all the information about the cross section of firms relevant for the aggregate dynamics in the model. As a result, the real side of the model aggregates to a model with a representative firm, where aggregate output is equal to

$$Y_t = K_t^{\phi} (e^{x_t} L_t)^{1-\phi}, \tag{16}$$

where the effective stock of capital K, defined in equation (7), evolves according to

$$\frac{dK_t}{K_t} = -\delta \, dt + \lambda \frac{e^{\xi_t}}{K_t} \left(\frac{I_t}{\lambda} \right)^{\alpha} dt, \tag{17}$$

where I_t is aggregate investment expenditures, which satisfy the aggregate resource constraint (12). The capital accumulation equation (17) illustrates that the embodied shock ξ acts as an investment-specific shock.

The net present value of new projects ν_t summarizes the marginal value of new investments. Specifically, the first-order condition for investment in equation (10), combined with market clearing, implies that in equilibrium,

$$I_t = \frac{\lambda \alpha}{1 - \alpha} \nu_t. \tag{18}$$

Examining equation (18), we see that conceptually ν_t plays a similar role in our model as Tobin's Q in a neoclassical model. However, there are two key differences. First, in our setting, the market value of a new project is not directly linked to the average Q of the firm. Second, equation (18) holds in levels, not ratios, as in the neoclassical model.

Aggregate quantities in the model (denominated in units of consumption) have both a permanent and a temporary component. To see this,

note that we can express most quantities of interest as functions of two aggregate state variables: a random walk component,

$$\chi_t = \frac{1-\phi}{1-\alpha\phi}\, x_t + \frac{\phi}{1-\alpha\phi}\,\xi_t, \tag{19}$$

and a stationary component,

$$\omega_t = \xi_t + \alpha\chi_t - \log K_t. \tag{20}$$

The trend variable χ_t is a function of the two technology shocks and thus determines long-run growth in the model. The variable ω determines the conditional growth rate in the effective capital stock K in equilibrium and therefore expected consumption growth. In a nonstochastic model, ω would be constant; in our stochastic model, ω is stationary. As a result, we can write aggregate output (16) as

$$\log Y_t = \chi_t - \phi\,\omega_t, \tag{21}$$

where we have used the fact that aggregate labor supply is constant, $L_t = 1$.

Aggregate consumption C, investment I, labor income w, and asset prices are cointegrated with aggregate output Y—they share the same stochastic trend, χ_t. Their transitory deviations from output (or χ_t) are functions of ω_t, which can be interpreted as the deviation of the current capital stock from its target level. Put differently, the state variable ω summarizes the transitory fluctuations of the model variables around the stochastic trend χ_t.

III. Illustrating the Model's Mechanism

To obtain some intuition about the asset pricing predictions of the model, we begin by analyzing the relation between technological progress, the SDF, and asset returns.

A. *The Pricing of Technology Risk*

The two technology shocks x and ξ affect the equilibrium SDF through their impact on the consumption of individual agents. Because of imperfect risk sharing, there is a distinction in how these shocks affect aggregate quantities and how they affect individual households. To emphasize this distinction, we first analyze the impact of technology on aggregate economic output and consumption and then examine its impact on the distribution of consumption for individual households.

1. Aggregate Quantities and Asset Prices

We calculate the impulse responses to the technology shocks x and ξ for aggregate output Y_t, consumption C_t, investment I_t, labor income w_t, and

aggregate payout to shareholders D_t, which is equal to total firm profits minus investment expenditures and payout to new inventors,

$$D_t = \phi Y_t - I_t - \eta \lambda \nu_t. \tag{22}$$

In the model, corporate payout is not restricted to be positive; nevertheless, using our parameter estimates, D becomes negative only in the extreme ranges of the state space that are not reached in model simulations.

Figure 2 plots these impulse responses.[9] Panel A shows that a positive disembodied technology shock x leads to an increase in output, consumption, investment, payout, and labor income. The increase in investment leads to higher capital accumulation, so the increase in output is persistent. However, since x is complementary to existing capital, most of its benefits are immediately realized. In panel B, we plot the response of these equilibrium quantities to a technology shock ξ that is embodied in new capital. In contrast to the disembodied shock x, the technology shock ξ affects output only through the formation of new capital stock. Consequently, it has no immediate effect on output and leads to only a reallocation of resources from consumption to investment on impact. Further, shareholder payout declines immediately after the shock, as firms cut dividends to fund new investments. In the medium run, the increase in investment leads to a gradual increase in output, consumption, payout, and the equilibrium wage (or labor income).

Next, we examine the impact of technology on the prices of financial assets and human capital. The total wealth of all existing households,

$$W_t \equiv \int_0^1 W_{n,t} dn = V_t + G_t + H_t, \tag{23}$$

is the sum of three components: the value of a claim on the profits of all existing projects $\mathcal{J}_t$,

$$V_t \equiv \int_0^1 \mathrm{E}_t \left[\sum_{j \in \mathcal{J}_{f,t}} \int_t^\infty \frac{\Lambda_s}{\Lambda_t} \Pi_{j,s} \right] df, \tag{24}$$

the value of new growth opportunities that accrues to shareholders,

$$G_t \equiv (1 - \eta) \int_0^1 \mathrm{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} \lambda_{f,s} \nu_s ds \right] df, \tag{25}$$

[9] To create plots, we use the parameter values in col. 3 of table A.1 in the appendix, which correspond to the baseline model discussed here. Since the goal of this section is to describe the qualitative features of the model, we postpone the discussion of how these parameters are estimated until sec. IV. That said, the model's qualitative implications are similar across several parametrizations.

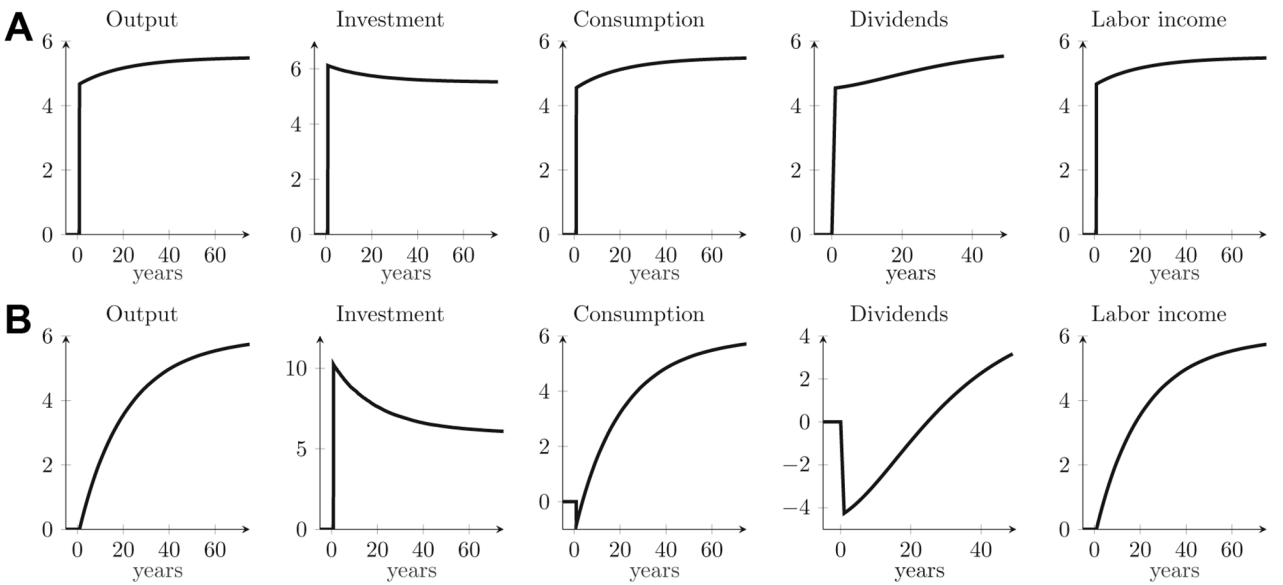


FIG. 2.—Technology and aggregate quantities. We plot the impulse response of aggregate output, investment, and consumption expenditures to the two technology shocks in the model. *A*, Response to x (disembodied shock). *B*, Response to ξ (embodied shock). We construct the impulse responses taking into account the nonlinear nature of equilibrium dynamics: we introduce an additional shock of 1 standard deviation at time $t = 0$ without altering the realizations of all future shocks. In our model, the scalar state variable ω summarizes all relevant information for the model's nonlinear dynamics. The impulse responses are computed at the mean of the stationary distribution of ω. We report the log difference between the mean response of the perturbed and unperturbed series (multiplied by 100).

and the present value of labor services to households,

$$H_t \equiv \mathrm{E}_t\left[\int_t^\infty e^{-\delta^h(s-t)}\,\frac{\Lambda_s}{\Lambda_t}\,\mathrm{w}_s ds\right]. \tag{26}$$

Figure 3 plots impulse responses for these components along with the risk-free rate and ν_t/W_t, which plays a key role in our subsequent analysis. A positive technology shock increases expected consumption growth, thereby increasing the risk-free rate. A positive disembodied shock x is complementary to installed capital, so the value of assets in place and growth opportunities rises. By contrast, an embodied shock ξ lowers the value of existing assets V while increasing the value of growth opportunities G. The net effect of an embodied shock is to lower total financial wealth $V + G$. By contrast, a disembodied shock leads to an increase in financial wealth.

The value of new blueprints ν relative to total wealth W rises in response to either technology shock, as we see in the last column of figure 3. Importantly, the effect is quantitatively much larger for the technological shock that is embodied in new projects compared with advances in technology that affect both existing and new projects. This difference is important because the responses of aggregate consumption in figure 2 to technology shocks mask substantial heterogeneity in the consumption paths of individual households. As we show next, larger changes in ν/W lead to greater reallocation of wealth among households.

2. Technology Shocks and Individual Consumption

The current state of a household can be summarized by its current share of total wealth, $w_{n,t} \equiv W_{n,t}/W_t$. The functional form of preferences (13) and (14) together with our assumption that the scale of the household-level innovation process is proportional to individual wealth imply that optimal individual consumption is proportional to individual wealth. Then, a household's consumption share is the same as its wealth share, $c_{i,t} = w_{i,t}$, and individual consumption satisfies $C_{i,t} = C_t\, w_{i,t}$. The dynamic evolution of households' share of aggregate wealth is

$$\frac{dw_{i,t}}{w_{i,t}} = \delta^h dt + \frac{\lambda}{\mu_I}\,\frac{\eta\,\nu_t}{W_t}\,(dN^I_{i,t} - \mu_I dt), \tag{27}$$

where $N^I_{i,t}$ is a Poisson process that counts the number of times that household i acquires a new blueprint.

The evolution of a household's relative wealth in equation (27) is conditional on the household's survival; thus, the first term captures the flow payoff of the annuity, as it is standard in perpetual youth overlapping generations models (Blanchard 1985). The second term captures changes in

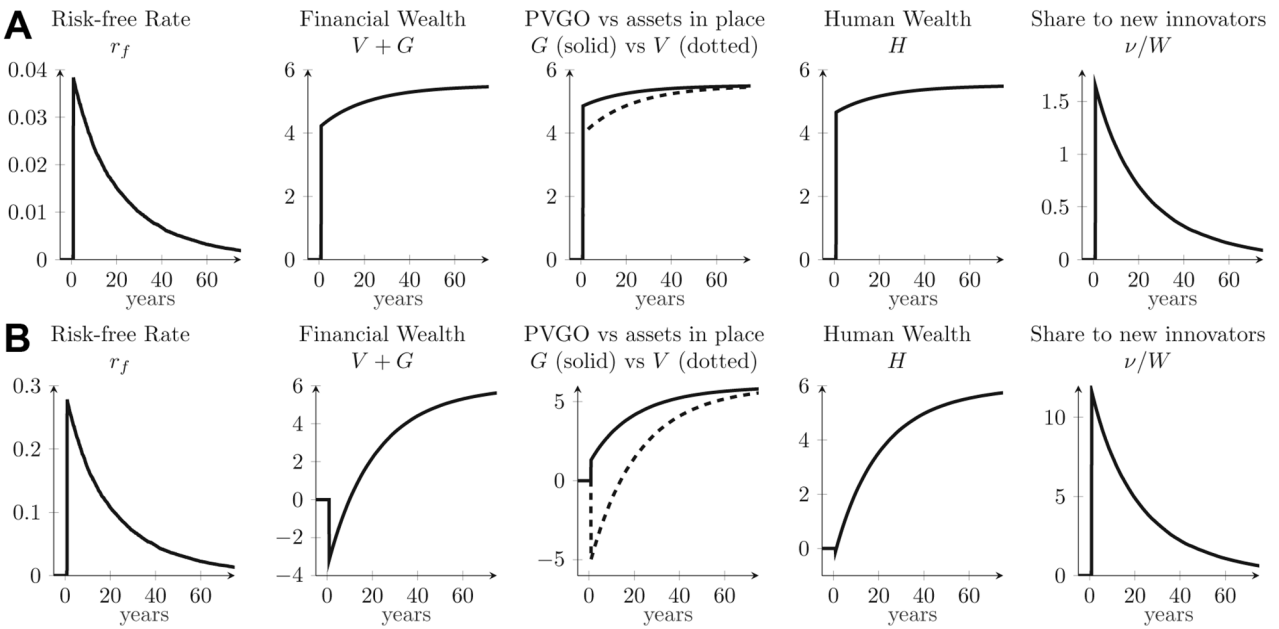


FIG. 3.—Technology and asset prices. We plot the impulse response of the risk-free rate (r_f) and the value of assets in place (V), growth opportunities (G), and human capital (H) as well as the relative value of payments to new innovators ν/W to the two technology shocks in the model. *A*, Response to x (disembodied shock). *B*, Response to ξ (embodied shock). We construct the impulse responses taking into account the nonlinear nature of equilibrium dynamics: we introduce an additional shock of 1 standard deviation at time $t = 0$ without altering the realizations of all future shocks. The impulse responses are computed at the mean of the stationary distribution of ω. In the first column, we report the difference between the mean response of the perturbed and unperturbed risk-free rate (multiplied by 100). For all other series, we report the log of the ratio of the perturbed to the unperturbed series (multiplied by 100).

the households' wealth resulting from innovation. Both the drift and the return to successful innovation depend on the fraction of shareholder wealth that accrues to all successful inventors, $\eta \nu_t / W_t$. This ratio is a monotonically increasing function of ω, defined in equation (20). Each period, a household yields a fraction $\lambda \eta \nu_t / W_t$ of its wealth share to successful innovators. This wealth reallocation occurs because households own shares in all firms, which make payments to new inventors in return for blueprints. During each infinitesimal time period, with probability $\mu_I \, dt$, the household is itself one of the innovators, in which case it receives a payoff proportional to $\eta \nu_t W_{i,t}$. The magnitude of wealth reallocation depends on the contribution of new investments to total wealth ν_t / W_t. As we saw in section III.A, an increase in either x or ξ implies that the equilibrium value of new blueprints to total wealth ν_t / W_t increases. This increase in ν_t / W_t leads to an increase in the households' idiosyncratic risk, similar to the model of Constantinides and Duffie (1996).

The process of wealth reallocation following positive technology shocks is highly skewed. Equation (27) shows that the rise in gains from successful innovation—that is, the rise in ν_t / W_t—implies that most households experience higher rates of relative wealth decline, as captured by a reduction in the drift of dw_t. By contrast, the few households that innovate increase their wealth shares greatly, as captured by the jump term $\nu_t / W_t \, dN_t^I$. From the perspective of a household at time t, the distribution of future consumption becomes more variable and skewed following a positive technological shock, even though on average the effect is zero—positive technological shocks magnify both extremely high realizations of w and the paths along which w declines persistently.

Figure 4 illustrates the impact of technology shocks on the consumption path of an individual household. Our objects of interest are the households' relative wealth share w_t, consumption C_t, and consumption adjusted for relative preferences $C_t^{1-h} w_t^h$. In the first three columns, we plot the impulse response of these variables to the two technology shocks. In addition to the response of the mean, we plot how the median of the future distribution of these variables changes in response to the shocks. Unlike the mean, the median of w is not influenced by rare but extremely positive outcomes and instead reflects the higher likelihood of large gradual relative wealth declines in response to technology shocks.

The first column of figure 4 summarizes the role of incomplete risk sharing in our model. Neither technology shock has an impact on the expected future wealth share w at any horizon, because in our model technology shocks have ex ante a symmetric effect on all households. However, the lack of an effect on the average wealth share masks substantial heterogeneity in individual outcomes. Specifically, the response of the median of the distribution is significantly negative at all horizons. The very different responses of the mean and the median wealth share

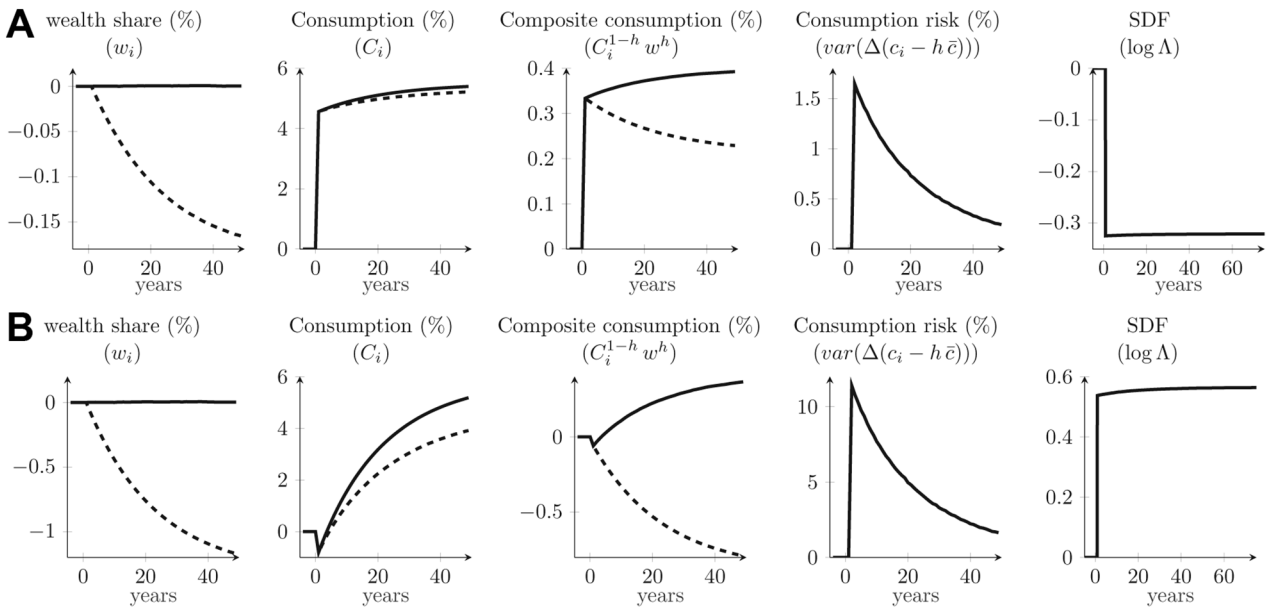


FIG. 4.—Technology, household consumption, and the stochastic discount factor. The first three columns plot the impulse response of the household wealth share (w_i), household consumption (C_i), and consumption adjusted for relative consumption preferences $C_i^{1-h} w_i^h$ to the two technology shocks in the model. *A*, Response to x (disembodied shock). *B*, Response to ξ (embodied shock). We construct the impulse responses taking into account the nonlinear nature of equilibrium dynamics: we introduce an additional shock of 1 standard deviation at time $t = 0$ without altering the realizations of all future shocks. The impulse responses are computed at the mean of the stationary distribution of ω. We report the log difference between the mean (median) response of the perturbed and unperturbed series with the solid (dashed) line. The fourth column plots the impulse response of the conditional variance of instantaneous log consumption growth, adjusted for relative preferences. The last column shows the impulse response of the equilibrium stochastic discount factor.

suggest a highly skewed effect of technological shocks on individual households, which is key to understanding the effects of technology shocks on the SDF in our model.

In the second column, we see that the responses of the mean and median future consumption to a disembodied shock x are not substantially different. A positive embodied shock ξ, however, leads to an increase in the share of value due to new blueprints ν_t/W_t and thus to greater wealth reallocation among households. The next column shows the role of relative consumption preferences. If households care about their relative consumption w_i in addition to their own consumption C_i, then the impact of technology shocks on their adjusted consumption is a weighted average of the first two columns.

The difference between the response of the mean and the median of the distribution of future consumption highlights the asymmetric benefits of technology shocks. Since households are risk averse, the mean response is insufficient to characterize the impact of technology on their indirect utility. When evaluating their future utility, households place little weight on the extremely high paths of w. Hence, the median response is also informative. In other words, in addition to their effect on the mean consumption growth, technology shocks also affect the variability of consumption because they affect the magnitude of the jump term in equation (27). Even though the conditional risk of individual wealth shares is idiosyncratic, this risk depends on the aggregate state of the economy. Therefore, innovation risk affects the SDF, similarly to the model of Constantinides and Duffie (1996). To illustrate this connection, the fourth column of figure 4 plots the increase in the variance of instantaneous consumption growth adjusted for relative preferences. We see that both technology shocks lead to higher consumption volatility. The effect is substantially higher for ξ than for x, again because of its higher impact on the returns to innovation ν_t/W_t.

In sum, figure 4 shows that the economic growth that results from technological improvements is not shared equally across households. Specifically, innovation reallocates wealth shares from most households to a select few. Although an increase in ν_t/W_t does not affect the expected wealth share of any household, it raises the magnitude of unexpected changes in households' wealth shares. Since households are risk averse, they dislike the resulting variability of changes in their wealth. Importantly, preferences over relative consumption ($h > 0$) magnify the negative effect of relative wealth shocks on indirect utility. Households are averse to displacement risk not only because it exposes them to additional consumption risk but also because they fear being left behind by subsequent increases in economic growth. Indeed, as we can see in the second column of panel B, a positive embodied shock increases the median consumption growth;

however, the third column shows that, once relative preferences are taken into account, the impact on the consumption bundle that incorporates relative consumption preferences is negative. The resulting effect on households' indirect utility has important implications about the pricing of these shocks, as we discuss next.

Last, we examine the SDF. Financial markets in our model are incomplete, since some of the shocks (specifically, the acquisition of blueprints by individual households) are not spanned by the set of traded financial assets. As a result, there does not exist a unique SDF in our model. Similar to Constantinides and Duffie (1996), the utility gradients of various agents are not identical, and each can serve as a valid SDF. To facilitate the discussion of the aggregate prices, we construct an SDF that is adapted to the market filtration $\mathcal{F}$ generated by the aggregate productivity shocks (B_x, B_ξ). This SDF is a projection of agent-specific SDFs (utility gradients) on $\mathcal{F}$. For more details, see section E in the appendix.

The risk prices of the two technology shocks can be recovered by the impulse response of the log SDF on impact. We plot these responses in the last column of figure 4. Comparing panels A and B, we see that the two technology shocks carry opposite prices of risk in our model. A positive disembodied shock x negatively affects the SDF on impact, implying a positive risk premium. By contrast, ξ has a negative risk premium. As a result, households value securities that provide a hedge against states where ξ is high and x is low.

The difference in how the SDF responds to the two technology shocks stems primarily from the response of the indirect utility term $f(\omega_t)$ in the SDF. Both technology shocks x and ξ lead to an increase in the permanent component χ_t of consumption, which causes the SDF to fall. However, in the case of the embodied shock, the fall in indirect utility due to the unequal sharing of benefits from technological progress is sufficiently large to offset the benefits of higher aggregate consumption. The resulting demand for insurance against high realizations of ξ is driven by the endogenous increase in the consumption uncertainty of individual investors.

B. Technology Shocks and the Cross Section of Firms

The firm's current state is fully characterized by the aggregate state X_t; the probability of acquiring new projects $\lambda_{f,t}$; the firm's relative size,

$$k_{f,t} \equiv \frac{K_{f,t}}{K_t}, \quad \text{where}$$
$$K_{f,t} \equiv \sum_{j \in \mathcal{J}_{f,t}} e^{\xi_{\tau(j)}} K_{j,t}; \tag{28}$$

and the firm's average profitability across projects,

$$\bar{u}_{f,t} \equiv \frac{1}{K_{f,t}} \left(\sum_{j \in \mathcal{J}_{f,t}} e^{\xi_{z(j)}} u_{j,t} K_{j,t} \right). \tag{29}$$

The assumption that shocks to projects' productivity u_j are perfectly correlated if they are owned by the same firm implies considerable dispersion in average profitability (29) across firms.

Our focus is on understanding differences in asset returns between value and growth firms. Typically, these categories are defined based on firms' valuation ratios, such as the firm's average Tobin's Q, or market-to-book ratio. In our model, a firm's Tobin's Q can be written as

$$\begin{aligned} \log Q_{f,t} - \log Q_t = \log \Bigg\{ & \frac{V_t}{V_t + G_t} \left[1 + \frac{v_1(\omega_t)}{v(\omega_t)} \left(\bar{u}_{ft} - 1 \right) \right] \\ & + \frac{G_t}{V_t + G_t} \frac{1}{k_{f,t}} \left[1 + \frac{g_1(\omega_t)}{g(\omega_t)} \left(\frac{\lambda_{f,t}}{\lambda} - 1 \right) \right] \Bigg\}, \end{aligned} \tag{30}$$

where Q_t is the market-to-book ratio of the market portfolio; V_t and G_t are defined in equations (24) and (25), respectively; and $v_1(\omega)$ and $g_1(\omega)$ are defined in lemmas 3 and 4 in the appendix.

Cross-sectional differences in valuation ratios (30) are mainly driven by the firm's current likelihood of future growth $\lambda_{f,t}$ relative to its current size $k_{f,t}$. Quantitatively, differences in $\bar{u}_f$ play only a minor role, as we demonstrate in the appendix. Firms with relatively high levels of λ_f / k_f are growth firms, since they derive more of their value from their future growth opportunities rather than their existing operations. Conversely, firms with low levels of λ_f / k_f derive most of their market value from their existing operations, and we therefore term them as value firms.

The impact of technology shocks on firm outcomes is related to the differences in their valuation ratios. Specifically, how firms respond to technology shocks depends on whether they are value or growth firms. Technological progress lowers the cost of new investments, hence it benefits firms with new investment opportunities (high $\lambda_{f,t}$). However, it also leads to displacement of installed capital because of general equilibrium effects and therefore harms firms with a lot of installed assets ($k_{f,t}$). To see these heterogeneous effects, we can express the profit flow of a firm f as

$$\Pi_{f,t} \equiv \sum_{j \in \mathcal{J}_{f,t}} \Pi_{j,t} = \phi \, Y_t \, \bar{u}_{f,t} \, k_{f,t}. \tag{31}$$

Equation (31) implies that cross-sectional differences in the sensitivity of firm profits to aggregate technology shocks can be summarized by how the firm's relative size k_f responds to shocks. The dynamics of k_f are given by

$$\frac{dk_{f,t}}{k_{f,t}} = a_0 \frac{\nu_t}{V_t}\left(\frac{\lambda_{f,t}}{\lambda k_{f,t}} - 1\right) dt + a_0 \frac{\nu_t}{V_t} \frac{1}{\lambda k_{f,t}} (dN_{f,t} - \lambda_{f,t} dt), \qquad (32)$$

where a_0 is a constant and $N_{f,t}$ is a Poisson process that counts the number of times that firm f has acquired a new project.

Just as technology shocks lead to wealth reallocation among households, they also lead to resource reallocation among firms. Equation (32) shows that the value of new blueprints ν relative to the value of installed capital V affects both the conditional growth rate as well as the dispersion in growth rates across firms. The ratio ν/V is a monotonically increasing function of the state variable ω. Focusing on the first term in (32), we see that high values of ν/V magnify the difference in conditional firm growth rates between growth firms (high levels of λ_f/k_f) and value firms (low levels of λ_f/k_f). The second term in equation (32) shows that high values of ν/V also increase the cross-sectional dispersion in firm growth rates. Recall that the ratio ν/V responds sharply to a positive embodied shock (ξ) but only modestly to a positive disembodied shock (x). Thus, technological innovation, especially when it is embodied in new capital, increases the rate of reallocation across firms, that is, the process of creative destruction.

To illustrate the heterogeneous impact of technology shocks on firm profitability and investment, we examine separately two firms with high and low levels of λ_f/k_f in figure 5. As we see in panel A, improvements in technology that are complementary to all capital lead to an immediate increase in profitability for both types of firms. Growth firms are more likely to have higher investment opportunities than the value firms; hence, on average, they increase investment. While growth firms pay lower dividends in the short run, their payouts rise over the long run. As a result, the market value of a growth firm appreciates more than the market value of a value firm. In panel B, we see that value and growth firms have very different responses to technology improvements embodied in new vintages. The technology shock ξ leaves the output of existing projects unaffected since it increases the productivity of only new investments. Because of the equilibrium response of the price of labor services, the profit flow from existing operations falls. Growth firms increase investment and experience an increase in profits and market valuations. In contrast, value firms have few new projects to invest in and therefore experience a decline in their profits and valuations.

In sum, growth firms have higher cash flow and stock return exposure to either technology shock than value firms, and the difference is quantitatively larger for technological improvements embodied in new capital vintages ξ versus shocks to labor productivity x. These differential responses of growth and value firms to technology shocks translate into cross-sectional differences in risk and risk premia.

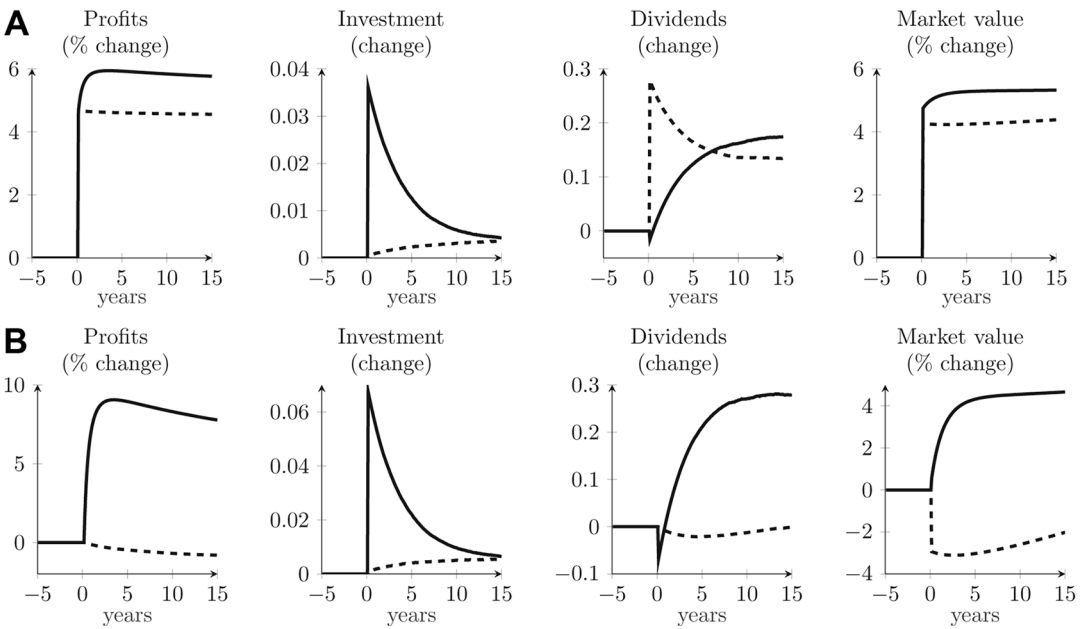


FIG. 5.—Technology shocks and firms. We plot the dynamic response of firm profits, investment, dividends, and stock prices to the two technology shocks in the model. *A*, Response to x (disembodied shock). *B*, Response to ξ (embodied shock). We construct the impulse responses taking into account the nonlinear nature of equilibrium dynamics: we introduce an additional shock of 1 standard deviation at time $t = 0$ without altering the realizations of all future shocks. We report separate results for two types of firms. The solid line represents the responses for a growth firm, defined as a firm with $\lambda_{f,t} = \lambda_H$ and $K_{f,t} = 0.5$. The dotted line indicates the responses for a value firm, defined as a firm with low investment opportunities $\lambda_{f,t} = \lambda_L$ and large size $K_{f,t} = 2$. For both firms, the level of average profitability is equal to its long-run mean, $u_{f,t} = 1$. The initial value of the state variable ω is set to its unconditional mean, $\omega_0 = E[\omega_t]$. Columns 1 and 4 plot percentage changes, and cols. 2 and 3 plot changes in the level (since both dividends and investment need not be positive) normalized by the aggregate dividend and investment at time $t = 0$.

IV. Estimation

In this section, we describe how we calibrate the model to the data. Given that the majority of households do not participate in financial markets, we first extend the model to include limited participation.

A. *Limited Participation*

A relatively small subset of households participate in financial markets. For instance, Poterba and Samwick (1995) report that the households in the top 20% in terms of asset ownership consistently own more than 98% of all stocks. Given the important role that inequality plays in the model, allowing for this type of limited participation is likely quantitatively important.

We model limited participation by assuming that newly born households are randomly assigned to one of two types, shareholders (with probability q_S) and workers (with probability $1 - q_S$). Shareholders have access to financial markets and optimize their lifetime utility of consumption, just like the households in our baseline model. Workers are instead hand-to-mouth consumers. They do not participate in financial markets, supply labor inelastically, and consume their labor income as it arrives. Workers can also successfully innovate (just like shareholders); those that do so become shareholders. Hence,

$$\psi \equiv \frac{\mu_I + q_S \delta^h}{\mu_I + \delta^h} \tag{33}$$

represents the fraction of households that participate in financial markets.

B. *Data and Estimation Strategy*

The market value of new blueprints ν_t plays a key role in the model's predictions both for the dynamics of firm cash flows (eq. [32]) and for the evolution of investors' wealth (eq. [27]). To take the model to the data, we use data on patents and stock returns to construct an empirical proxy for ν_t, following Kogan et al. (2017). To examine the implications of the model, we aggregate these patent values by constructing estimates of the average value of innovations in a given year ν_t at both the aggregate as well as the firm level. To preserve stationarity, we scale these by (aggregate or firm) market capitalization. We refer to these two measures as aggregate and firm innovation, respectively. Section D in the appendix provides further details on the construction of the variables used in estimating the model.

The model has a total of 21 parameters. We choose the probability of household death as $\delta^h = 1/40$ to guarantee an average working life

of 40 years. We choose the likelihood of repeat innovation $\mu_I = 0.13\%$ so that, conditional on the other parameters, the model's average top 1% income shares are in line with the data. We estimate the remaining parameters of the model using a simulated minimum distance method (Ingram and Lee 1991). Specifically, given a vector X of target statistics in the data, we obtain parameter estimates by

$$\hat{p} = \arg\min_{p \in \mathcal{P}} \left(X - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(p) \right)' W \left(X - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(p) \right), \tag{34}$$

where $\hat{X}_i(p)$ is the vector of statistics computed in one simulation of the model. Our choice of weighting matrix $W = \text{diag}(XX')^{-1}$ penalizes proportional deviations of the model statistics from their empirical counterparts.

Our estimation targets are reported in the first column of table 1. They include a combination of first and second moments of aggregate quantities, asset returns, and firm-specific moments. Many of the moments that we target are relatively standard in the literature. Others are less common, but they are revealing of the main mechanisms of our paper. First, both the investment-to-output ratio (I_t/Y_t) as well as the relative value of new innovations (ν_t/M_t) are increasing functions of ω in the model. Thus, the unconditional volatility of these ratios is informative about the model parameters. Similarly, fluctuations in ω lead to predictable movements in consumption growth. We therefore also include as a target an estimate of the long-run volatility of consumption growth using the methodology of Dew-Becker (2017) in addition to its short-term (annual) volatility.

Second, our model connects embodied technology shocks to the return differential between value and growth firms. We thus include as estimation targets not only the first two moments of the market portfolio but also the average value premium, defined as the difference in risk premia between firms in the bottom versus top decile in terms of their market-to-book ratios (Q), following Fama and French (1992). Given that the model has no debt, we create returns to equity by levering the value of corresponding dividend (payout) claims by a factor of 2.5. This value lies between the estimates of the financial leverage of the corporate sector of Rauh and Sufi (2011; which is equal to 2) and the values used by Abel (1999) and Bansal and Yaron (2004; 2.7–3).

C. Model Fit

Table 1 shows that the baseline model fits the data reasonably well. The model not only generates realistic patterns for aggregate consumption and investment but also can fit both the mean as well as the dispersion in risk premia in the cross section of firms. In addition, the model generates

TABLE 1
BENCHMARK MODEL: GOODNESS OF FIT

		Model			
Statistic	Data	Mean	5th	95th	SQRD
Aggregate quantities:					
Consumption growth, mean	.015	.014	.003	.024	.003
Consumption growth, volatility (short run)	.036	.038	.035	.043	.005
Consumption growth, volatility (long run)	.041	.053	.039	.068	.083
Shareholder consumption share, mean	.429	.464	.437	.491	.006
Shareholder consumption growth, volatility	.037	.039	.025	.053	.002
Investment-to-output ratio, mean	.089	.083	.043	.123	.005
Investment-to-output ratio (log), volatility	.305	.288	.149	.506	.003
Investment growth, volatility	.130	.105	.083	.126	.037
Investment and consumption growth, correlation	.472	.373	.170	.537	.044
Aggregate innovation, volatility	.370	.369	.219	.619	.000
Capital share	.356	.354	.312	.385	.001
Asset prices:					
Market portfolio, excess returns, mean	.063	.067	.052	.080	.003
Market portfolio, excess returns, volatility	.185	.131	.119	.143	.087
Risk-free rate, mean	.020	.020	.015	.029	.003
Risk-free rate, volatility	.007	.007	.003	.013	.000
Value factor, mean	.065	.063	.031	.093	.001
Cross-sectional moments:					
Investment rate, interquartile range	.175	.163	.124	.204	.005
Investment rate, persistence	.223	.228	.053	.457	.000
Tobin Q, interquartile range	1.139	.882	.611	1.147	.051
Tobin Q, persistence	.889	.948	.928	.961	.004
Firm innovation, 90–50 range	.581	.542	.441	.609	.005
Firm innovation, persistence	.551	.567	.519	.623	.001
Profitabiltiy, interquartile range	.902	.936	.856	1.073	.001
Profitability, persistence	.818	.815	.796	.830	.000
Distance (mean squared relative deviation)					.014

NOTE.—This table reports the fit of the model to the statistics of the data that we target. Growth rates and rates of return are reported at annual frequencies. See main text for details on the estimation method and sec. B in the appendix (available online) for details on the data construction. We report the mean statistic along with the 5th and 95th percentiles across simulations. We also report the squared relative deviation of the mean statistic to their empirical counterparts, $\text{SQRD}_i = (X_i - \hat{X}_i(p))^2 / X_i^2$.

realistic levels of dispersion and persistence in firm-level profitability, investment, innovation output, and valuation ratios.

On the asset pricing side, the main success of our model relative to existing work is that it delivers realistic predictions about not only the equity premium but also the cross section of asset returns. Specifically, the model generates realistic differences in risk premia between high-Q (growth) and low-Q (value) firms. These patterns arise primarily from the fact that the two technology shocks in the model carry opposite risk prices (as we see in the last column of fig. 4) and the fact that value and growth firms have differential exposures to the two technology shocks (as we see in fig. 5). Further, even though these are not our main objects of interest,

the model also generates a low and stable risk-free rate. In this respect, our model represents an improvement compared with many equilibrium asset pricing models with production (see, e.g., Jermann 1998; Kaltenbrunner and Lochstoer 2010).[10] The model does have some difficulty in generating sufficiently volatile stock returns—stock returns in the model are smoother than their empirical counterparts by approximately 30%. As is the case with almost all general equilibrium models, the need to match the relatively smooth dynamics of aggregate quantities imposes tight constraints on the shock volatilities σ_x and σ_ξ.

The model's success in fitting the asset pricing moments does not come at the cost of counterfactual implications for quantities. Examining the first ten rows of table 1, we see that the model generates realistic implications for aggregate quantities. Even in the few cases in which the point estimates between the model and the data differ, the empirical moments can still be plausibly observed in simulated data; that is, they are covered by the standard 90% confidence intervals based on model simulations. Similarly, the last nine rows of table 1 show that the model also generates realistic firm-level moments. This is notable in itself since existing models that fit the cross section of asset returns fail to match the cross section of firm investment, and vice versa (see, e.g., Clementi and Palazzo 2019). Further, a notable feature of the data is that firm-level investment and innovation exhibit relatively low persistence, while valuation ratios (Q) are highly persistent. The model can largely accommodate this behavior because realized investment at the firm-level exhibits spikes; that is, firms only invest conditional on having a project. Valuation ratios depend partly on expectations of future investment and therefore are much more persistent.

D. Parameter Estimates and Identification

Table 2 reports the parameter estimates in the model along with standard errors. To obtain some intuition for which features of the data identify individual parameters, we compute the Gentzkow and Shapiro (2017) measure of (local) sensitivity of parameters to moments. To conserve space, we summarize the results only for the parameters related to the nonstandard features of the model. We relegate a more comprehensive discussion of identification to section A in the appendix.

The parameters governing the volatility of the two technology shocks are of primary importance for the model's implications about the dynamics

[10] These models are often hampered by the fact that consumption rises while dividends fall after a positive technology shock, leading to a negative correlation between aggregate payout of the corporate sector and consumption (see, e.g., Rouwenhorst 1995). In sec. C.4 in the appendix, we show that consumption and dividends are positively correlated in our setup, which helps the model deliver a sizeable equity premium.

TABLE 2
BENCHMARK MODEL: PARAMETER ESTIMATES

Parameter	Symbol	Estimate	Standard Error
Preferences:			
Risk aversion	γ	56.734	41.163
Elasticity of intertemporal substitution	θ	2.341	2.867
Effective discount rate	ρ	.044	.015
Preference weight on relative consumption	h	.836	.067
Technology:			
Disembodied technology growth, mean	μ_x	.016	.011
Disembodied technology growth, volatility	σ_x	.082	.010
Embodied technology growth, mean	μ_ξ	.004	.018
Embodied technology growth, volatility	σ_ξ	.110	.025
Project-specific productivity, volatility	σ_u	.533	.065
Project-specific productivity, mean reversion	κ_u	.210	.023
Production and investment:			
Cobb-Douglas capital share	ϕ	.427	.023
Decreasing returns to investment	α	.446	.096
Depreciation rate	δ	.029	.011
Transition rate to low-growth state	μ_L	.364	.098
Transition rate to high-growth state	μ_H	.021	.006
Project mean arrival rate, mean	λ	.812	.273
Project mean arrival rate, relative difference in high- and low-growth states	λ_D	15.674	3.303
Incomplete markets:			
Fraction of project net present value that goes to inventors	η	.767	.391
Fraction of households that is a shareholder	ψ	.148	.060

NOTE.—This table reports the estimated parameters of the model. When constructing standard errors, we approximate the gradient of $(1/S)\Sigma_{i=1}^{S}\hat{X}_i(p)$ using a five-point stencil centered at the parameter vector $\hat{p}$. In addition, we approximate the sampling distribution of X (the empirical moment covariance matrix) with the sampling distribution of $X_i(p)$ in the model (the covariance matrix of model moments across simulations. See the appendix (available online) for more details.

of aggregate quantities and asset prices. Accordingly, their estimated values $\hat{\sigma}_x = 8.2\%$ and $\hat{\sigma}_\xi = 11\%$ are estimated with considerable precision. Examining the Gentzkow and Shapiro sensitivity measure, we see that σ_x is primarily identified by the volatility of consumption growth and the correlation between consumption and investment. Conversely, σ_ξ is mostly identified by the volatility of investment growth and the correlation between investment and consumption. Recall from figure 2 that both investment and consumption respond symmetrically to the disembodied shock x; by contrast, investment and consumption respond initially with opposite signs to ξ. Hence, the correlation between investment and consumption carries important information on the relative importance of these two shocks.

The parameter that governs the share of households that participate in the stock market ψ is relatively well estimated, with a point estimate of 0.15 and a standard error of 0.05. Our estimate is largely in line with the facts reported by Poterba and Samwick (1995), who document that the households

in the top 20% in terms of asset ownership consistently own more than 98% of all stocks. This parameter is primarily identified by the mean consumption share of stockholders. The parameter that affects the division of surplus between shareholders and innovators is estimated at $\eta = 0.77$ with a standard error of 0.39. Hence, approximately one-fourth of the value of new investment opportunities in the economy accrues to the owners of the firm's public securities.[11] In general, this parameter has an ambiguous effect on the value premium. On the one hand, increasing η increases the share of rents that go to innovators and therefore increases the displacement risk that is faced by shareholders. On the other hand, however, it reduces the overall share of growth opportunities to firm value, which increases the volatility of the market portfolio since the market is now more vulnerable to displacement. Consequently, this parameter is mainly identified by the average returns of the market portfolio and the value factor (which are noisily estimated) and, to a lesser extent, by the long-run volatility of consumption. A higher long-run volatility of consumption growth implies that the model needs a smaller value of η to match asset prices. Last, the parameter governing the share of relative consumption h is identified primarily by the differences in risk premia between the market and the value factor. Recalling the discussion in figure 4, higher values of h imply that households are more averse to displacement risk—they resent being left behind—and hence increase the risk premium associated with displacement risk. However, higher values of h also lower the household's effective risk aversion toward aggregate—that is, nondisplacive—shocks. Since the equity premium compensates investors partly for this aggregate risk, increasing h lowers the equity premium while increasing the mean returns of the value factor. Our baseline estimates imply that households attach a weight approximately equal to 80% on relative consumption. The small standard error (0.06) implies that this parameter is quantitatively important for the implications of the model.

E. Sensitivity Analysis

Our model has several relatively nonstandard features. To quantify how each of these features contributes to our findings, we estimate the model

[11] A quantitative evaluation of the plausibility of this parameter is challenging because of the lack of available data on valuations of private firms. However, if we interpret the selling of ideas to firms as the inventors creating new startups that go public and are subsequently sold to large, publicly traded firms, this pattern is consistent with the empirical fact that most of the rents from acquisitions go to the owners of the target firm (Asquith and Kim 1982). In addition, a high estimate of η is consistent with the idea that ideas are a scarcer resource than capital, and since the innovators own the ideas, it is conceivable that they extract most of the rents from the creation of new projects. Another real-world example of an inventor is a venture capitalist. A large value of η would be consistent with the evidence that venture capitalist firms add considerable value to startups, but investors in these funds effectively capture none of these rents (see, e.g., Korteweg and Nagel 2016).

imposing several parameter restrictions. We summarize our conclusions here and refer the reader to section B in the appendix for further details. In brief, we find that the following three features are key for the performance of the model: incomplete markets for innovation, preferences for relative consumption, and embodied technology shocks. Eliminating any of these three features compromises the model's ability to simultaneously fit both the level and the cross section in risk premia. In addition, the model can largely accommodate a strong prior belief about an upper bound on η or h, as long as these features are not eliminated fully. Put differently, cross-sectional differences in risk premia arise because of displacement risk. To fit the data, the model requires that there is sufficient displacement risk and that households care enough about displacement. By contrast, the assumption of limited participation is less quantitatively important: the model can still fit the data, albeit with a higher estimated coefficient of relative risk aversion. Last, restricting the coefficient of relative risk aversion to lie below 10 still results in risk premia that are not too different from the data.

V. Additional Predictions: Model versus Data

In this section we examine the performance of the model in replicating some features of the data that we do not use as explicit estimation targets, including income and wealth inequality, asset pricing anomalies of value and growth firms, correlations between consumption growth and asset returns or corporate payout, and the term structure of interest rates and risk.

A. Consumption, Wealth, and Income Inequality

One of the advantages of our incomplete markets model is that it has implications for inequality. In the model, inequality arises because households cannot share the rents to innovation. Lucky households that successfully innovate experience a temporary increase in their income because of the proceeds from innovation; in addition, their wealth (and consumption) also increases. Here, we compare the implications of the model for several measures of consumption, wealth, and income inequality in the data. Since the model is largely silent regarding inequality between middle- and low-income households, whenever possible, we focus on the top end of the distribution.

We compute the following measures of inequality. Using the data of Piketty and Saez (2003) and Saez and Zucman (2016), we compute average top income and wealth shares at the 0.1%, 0.5%, and 1% cutoffs. In addition, we report top percentile ratios in income and wealth (net worth) using the Survey of Consumer Finances (SCF) weights. In order to be consistent with the model, we report these statistics only for the households

that report direct or indirect stock ownership. We also remove cohort and age effects, although this does not have a major impact on these estimates. For completeness, we report corresponding statistics using the Consumer Expenditure Survey (CEX). However, these should be interpreted with caution because wealthy households are largely absent from the CEX. By contrast, the SCF oversamples wealthy households, so the moments for wealth and income inequality are likely to be more reliable. We replicate the same statistics using a long simulation from the model. To be as close to the data as possible, we use the income, or wealth, of all households in the denominator to make these estimates comparable with the data of Piketty and Saez (2003) and Saez and Zucman (2016). By contrast, when estimating the ratio of top percentiles, we use only the subset of households that participate in financial markets.

Table 3 presents the results. Recall that the parameter μ_I is chosen to match the income share of the top 0.1%. The rest of the statistics in table 3 are not part of the estimation targets. Examining the table, we note that the model is largely successful in replicating the magnitude of income and wealth inequality, especially at the very top. Specifically, the top income and wealth shares are largely in line with the data. Further, focusing on the

TABLE 3
INEQUALITY MOMENTS

	CONSUMPTION		WEALTH		INCOME	
	Model	Data	Model	Data	Model	Data
Top shares (all households; %):						
.1	5.3	.7	12.1	13.2	5.6	5.6
.5	11.2	2.5	26.2	26.2	11.5	10.8
1.0	16.7	4.2	36.5	32.6	15.6	14.2
Percentile ratios (stockholders):						
99–90	3.3	1.7	3.3	5.7	3.5	4.2
99–95	2.4	1.5	2.4	3.2	2.5	2.9
95–90	1.4	1.2	1.4	1.8	1.4	1.5
90–50	5.6	1.8	5.6	5.1	5.6	2.5

NOTE.—This table compares measures of inequality between the model and the data. Data sources are the CEX, the SCF, Piketty and Saez (2003), and Saez and Zucman (2016). The top income and wealth shares are from Piketty and Saez (2003) and Saez and Zucman (2016). Top consumption shares are from the CEX (1982–2010). Top shares are calculated relative to all households. The top percentile shares of income (total income) and wealth (net worth) are from the SCF (1989–2013); we report percentile ratios of the stock ownership sample (equity = 1 in the SCF summary extracts) and after obtaining residuals from cohort and year dummies and cubic age effects. We use a similar procedure for the CEX data. The corresponding estimates in the model are computed from a long simulation of 10 million households over 10,000 years. Percentile ratios are computed among the subset that participates in the stock market. In the model, income equals wages, payout, and proceeds from innovations. In the data, we use the total income variable from the SCF, which includes salary, proceeds from owning a business, and capital income. In the Piketty and Saez (2003) data, we use income shares inclusive of capital gains.

distribution of income and wealth among the stockholder sample, we can see that the model can largely replicate the empirical patterns.

The model differs from the data along some dimensions. First, the ratio of income between households in the 90th–50th percentile is much higher in the model than in the data. The reason is that in the model, households are free to retrade claims on their future labor services w_t; wealthy households therefore purchase a larger fraction of human wealth in order to have the same ratio of financial to human wealth. This feature is essential in preserving the analytic tractability of the model but also implies somewhat larger income inequality among stockholders in the middle of the distribution. Second, consumption inequality is somewhat higher in the model than in the CEX data. In the model, all shareholders consume the same fraction of their wealth, hence consumption inequality among stockholders is the same as wealth inequality. Top consumption shares differ from top wealth shares because of the presence of nonparticipating households, who simply consume their wage income. However, an important caveat in this comparison is that wealthy households are undersampled in the CEX, hence the empirical moments of consumption inequality should be interpreted with caution.

In sum, we see that the model generates realistic patterns of inequality, especially at the very top. This is important for several reasons. First, these moments indirectly rely on the parameters we estimated in section IV, for instance, the share of surplus that accrues to innovators η. The fact that the calibrated value produces realistic moments for inequality is comforting, given those parameters, despite the fact that these were not explicit targets for calibration (with the exception of the income share of the top 1% that is used to calibrate μ_1). Second, to the best of our knowledge, this is the first general equilibrium asset pricing model that can generate realistic patterns of inequality.

B. Cross-Sectional Asset Pricing Puzzles

A stylized feature of the asset returns is that the CAPM does a poor job fitting the cross section of risk premia (Fama and French 1992). A closely related puzzle is the existence of risk factors in the cross section of returns, for instance, the value factor. That is, spread portfolios formed by trading on the top and bottom deciles of characteristic sorted portfolios not only have sizable CAPM alphas (they are mispriced by the CAPM) but also are substantially volatile while also exhibiting low correlation with the market portfolio. As Cochrane (2005) writes, this comovement puzzle is a challenge for existing general equilibrium models. These models largely imply that the market portfolio is a summary statistic for all aggregate risk in the economy.

Our model can replicate these patterns, even though they are not part of the estimation targets. As we see in the top two rows of table 4, the model not only generates realistic differences in risk premia between high-Q (growth) and low-Q (value) firms but also generates return comovement—the value factor of Fama and French (1993). In both the model and the data, the value minus growth portfolio is highly volatile but essentially uncorrelated with the market portfolio. In addition, it replicates the failure of the CAPM in accounting for these return differences. These patterns arise primarily from the fact that the two technology shocks in the model carry opposite risk prices (as we see in the last column of fig. 4) and the fact that value and growth firms have different exposures to the two technology shocks (as we see in fig. 5). Specifically, as we can see in the bottom six rows of table A.2, the value factor and the market portfolio have different exposures to the disembodied shock x; since the disembodied shock x has to be rather volatile to match the volatility of consumption—and the correlation between consumption and investment—this pushes the model

TABLE 4
CROSS-SECTIONAL ASSET PRICING PUZZLES

	Mean	Standard Deviation	CAPM α	CAPM β	R^2
BE/ME spread:					
Model	.061	.165	.041	.254	.071
Data	.070	.216	.049	.262	.053
Standard error	.025	.004	.026	.038	.015
I/K spread:					
Model	−.020	.173	−.015	−.080	.009
Data	−.056	.112	−.067	.190	.068
Standard error	.017	.003	.016	.029	.023
E/P spread:					
Model	.040	.110	.020	.252	.099
Data	.060	.201	.068	−.131	.013
Standard error	.019	.004	.032	.170	.010

NOTE.—This table reports the mean returns, volatilities, CAPM α's and β's, and regression R^2 from a market model for three spread portfolios: the value spread, defined as the difference in returns between the bottom and the top decile in terms of market-to-book ratio; the investment spread (I/K), defined as the return spread in decile portfolios of firms sorted on the basis of their past investment rate; and the earnings-price spread (E/P), defined as the returns spread in decile portfolios on the basis of earnings-to-price ratios. Model estimates are based on a long simulation of the model (10,000 years). The moments in the data are based on portfolios available from Kenneth French (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). Data period for the value premium excludes data prior to the formation of the SEC (1936–2010); data period for the investment strategy (I/K) is 1964–2010; data period for the earnings-to-price strategy is 1952–2010. Standard errors for the empirical moments are included. Standard errors for R^2 are computed using the delta method.

to have a relatively low correlation between the market and the value factor, even when this correlation is not an explicit part of the estimation targets. As long as the model matches both the equity and the value premium, this low correlation between the market and value will result in a substantial CAPM alpha.

In the next set of rows, we show that the model can, to a lesser degree, also replicate two other closely related anomalies in the data, the high-minus-low investment strategy (Titman, Wei, and Xie 2004) and the high-minus-low earnings-to-price strategy (Rosenberg, Reid, and Lanstein 1985). The mechanism that delivers these patterns is essentially the same as that which delivers the value spread: firms with high investment and low earnings-to-price ratios are high growth firms (high λ_f/K_f) and are valued by investors because they help hedge displacement risk. However, at the current parametrization, the model cannot fully match the differences in risk premia observed in the data: the model-implied risk premia are approximately one-fourth to one-half of their empirical values. In particular, in the case of the investment spread, in our current model, investment is a noisy proxy for the firm's current growth opportunities. We conjecture that extending the production side of the model along the lines of Kogan and Papanikolaou (2013) will deliver a much closer fit to the data.

C. Consumption and Asset Returns

We next examine the implications of the model for the joint distribution of consumption and asset returns, specifically, returns to the market portfolio and the value factor. Since our baseline model features limited participation, we report results separately for stockholders, using the data of Malloy, Moskowitz, and Vissing-Jorgensen (2009). We compute correlations of asset returns and dividends from the market portfolio with the consumption of stockholders in absolute terms but also relative to the consumption of nonparticipants. We follow standard practice and aggregate consumption growth over multiple horizons. We report results using 2-year growth rates, but the results are qualitatively similar using longer horizons.

Table 5 shows that the baseline model generates empirically plausible levels of correlation between asset returns and aggregate consumption. Consumption and aggregate stock market returns are more highly correlated in the model, but the difference between the correlations in the model and in the data is not always statistically significant. The model also reproduces the low empirical correlation between aggregate consumption growth and the value factor, which implies the failure of the consumption CAPM to capture the value premium in simulated data. Further, we see that the model replicates one of the main findings of Malloy, Moskowitz, and Vissing-Jorgensen (2009): returns of value firms covary more with

TABLE 5
CONSUMPTION AND ASSET RETURNS

	AGGREGATE		SHAREHOLDER	
CONSUMPTION CORRELATION WITH	Market Portfolio	Value Factor	Market Portfolio	Value Factor
Model	.66	.00	.71	.12
Data	.41	.16	.21	.35
Standard error	.15	.13	.12	.18

NOTE.—This table compares the empirical correlations between three consumption measures (aggregate consumption, consumption of stockholders, and relative consumption of stockholders) with the corresponding correlations in the model. Market returns are in excess of the risk-free rate. The value factor is the difference in returns between stocks in the top and bottom deciles in terms of book-to-market ratio. The empirical correlations with shareholder consumption are based on the data of Malloy, Moskowitz, and Vissing-Jorgensen (2009), which cover the 1982–2002 period. We use the convention of computing correlations of 1-year asset returns with 2-year consumption growth.

shareholder consumption than returns of growth firms. The model output does differ from the data in one respect: it generates a somewhat excessive correlation of shareholder consumption and the market portfolio (equal to 21% in the 1982–2002 sample vs. 71% in the model).

D. Term Structure of Interest Rates

Next, we examine the predictions of our model for the price of zero coupon bonds. The price of a zero coupon bond of maturity T can be computed as $P_b(t, T) = E_t(\Lambda_T/\Lambda_t)$, while zero coupon yields can be computed as

$$y(T, t) = -\frac{1}{T}\log P_b(t, T). \tag{35}$$

We solve for equation (35) numerically and report the moments of bond yields in table 6. The model generates a real yield curve that is on average

TABLE 6
MODEL: TERM STRUCTURE MOMENTS

	YEARS									
MATURITY	1	2	3	4	5	6	7	8	9	10
Yield, average (%)	2.0	2.0	2.1	2.2	2.2	2.3	2.4	2.4	2.5	2.6
Yield, volatility (%)	.7	.7	.8	.8	.8	.8	.8	.8	.8	.9
Excess return, mean (%)		.1	.3	.5	.6	.8	.9	1.1	1.3	1.4
Excess return, volatility (%)		.3	.6	.9	1.2	1.5	1.8	1.8	2.3	2.5

NOTE.—This table reports the moments for the term structure of real interest rates that is implied by the model under the baseline calibration. Bond excess returns are reported in excess of the risk-free rate. We report moments across a long simulation of the model (10,000 years).

upward sloping. In addition, average excess returns on real bonds are positive and rise with maturity. Real bonds are risky in the model because the risk-free interest rate is positively correlated with the SDF. Specifically, recall that in figure 3 interest rate is mostly sensitive to the embodied shock ξ, which, as figure 4 shows, is positively correlated with the SDF.

The implications of our model for the yield curve are in stark contrast with most leading equilibrium asset pricing models, which typically predict a downward-sloping real yield curve and negative yields on real bonds (Bansal and Yaron 2004; Bansal, Kiku, and Yaron 2012; Wachter 2013).[12] By contrast, the observed term structure of Treasury inflation-protected securities (TIPS) has never had a significant negative slope, and the real yield on long-term TIPS has always been positive and is usually above 2% (Beeler and Campbell 2012).

E. Term Structure of Risk and Risk Prices

Our model generates a sizable equity premium due to joint movements in aggregate dividends and the SDF. Here, we briefly examine how risk in the total payout of the market portfolio at different horizons contributes to asset prices. To do so, we compute prices of corporate payout strips, that is, a claim on the total payout of the corporate sector at a future date T,

$$P_d(t, T) = E_t\left[\frac{\Lambda_T}{\Lambda_t} D_T\right], \tag{36}$$

where total payout D_t is given by equation (22).

In figure 6, we plot the risk premia and volatilities of corporate strips across maturities of 1–20 years. We report results for unlevered claims. In panel A, we see that average returns of dividend strips are decreasing sharply across maturities. Specifically, risk premia range from 4.8% for the shortest maturity strips (1 year) to 2.1% for strips with maturity of 20 years. For comparison, the mean returns on an unlevered claim to corporate payout is 2.9%. Hence, the equity premium in the model is concentrated at shorter maturities. In panel B, we also see that the prices of dividend strips at shorter maturities are more volatile than those of longer-maturity strips.

The above patterns are qualitatively consistent with recent empirical literature. Specifically, van Binsbergen, Brandt, and Koijen (2012) show that risk premia of dividend strips on the market portfolio exhibit a downward-sloping term structure, with the short end of the curve accounting for most of the equity premium. In addition, they show that shorter maturity

[12] A notable exception is Campbell and Cochrane (1999): the working paper version shows that their endowment economy model also replicates an upward-sloping real yield curve.

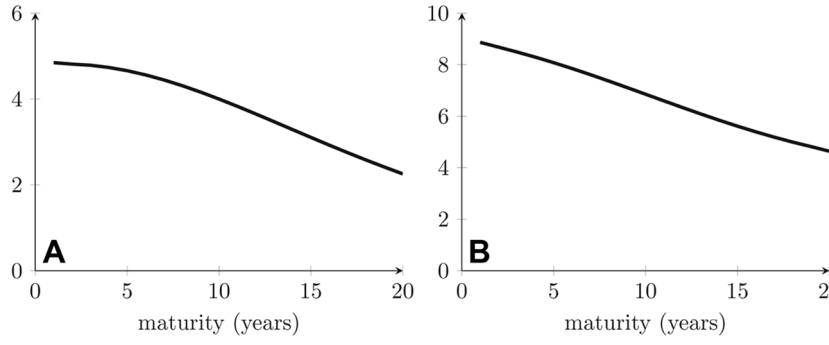


FIG. 6.—Model-implied moments of corporate payout strips. *A*, Mean excess returns (%). *B*, Return volatility (%). We plot the return means and standard deviations for the term structure of corporate strips that are implied by the model under the baseline calibration. We compute returns in excess of the risk-free rate and without leverage, using a long simulation of the model (10,000 years).

dividend claims are more volatile. Last, they argue that these patterns are in stark contrast with most existing equilibrium asset pricing models, although an important caveat in this comparison is that the empirical evidence pertains to claims on cash dividends, whereas the model results pertain to net payout.

In our model, these patterns arise naturally in equilibrium. Specifically, we saw in panel A of figure 2 that the contribution of the dividend dynamics induced by x to the equity premium rises modestly with the horizon. Panel B implies the opposite pattern; the contribution of the dividend dynamics induced by ξ to the equity premium is concentrated in the short and medium run, and the rise in long-run dividends contributes negatively to the equity premium. Thus, the term structure of dividend strip risk premia is downward sloping. To conserve space, we refer the reader to section C.1 in the appendix for more details.

F. Innovation and Aggregate Quantities

Our analysis thus far closely follows the existing literature and evaluates the success of the model on the basis of the model-implied correlations between macroeconomic quantities and prices. In the remainder of this section, we use the estimated value of new innovations that we constructed using the methodology of Kogan et al. (2017) to examine more directly the predictions of the model's main mechanism.

We begin by documenting the correlation between the rate of innovation ν/M and aggregate quantities and prices. We then compare the empirical results with those in simulated data. One potential shortcoming of our empirical measure of ν_t is that it identifies the timing of the

innovation with the year when that patent is issued to the firm. While this timing is helpful in estimating the value of the patent on the basis of the firm's stock market reaction, it is an imprecise measure of the actual timing of when the invention took place. To ensure that this potential timing mismatch does not affect our results, we report correlations across periods of 1–3 years.

In panels A and B of table 7, we compare the correlations between (log differences of) the rate of innovation ν_t/M_t and consumption growth—both aggregate consumption as well as the consumption of shareholders. We see that the rate of innovation has a weak negative relation with both measures of consumption growth. In our model, innovation is also

TABLE 7
INNOVATION, AGGREGATE QUANTITIES, AND ASSET RETURNS: MODEL VERSUS DATA

VALUE OF NEW INVENTIONS (ν_t/M_t)	CORRELATION		
	$t \to t+1$	$t \to t+2$	$t \to t+3$
	A. Consumption Growth, Aggregate		
Model	.04	.00	.00
Data	−.14	−.17	−.16
Standard error	.16	.16	.15
	B. Consumption Growth, Shareholders		
Model	−.13	−.18	−.18
Data	−.22	−.13	−.13
Standard error	.11	.11	.13
	C. Investment Growth		
Model	.72	.76	.78
Data	.30	.27	.30
Standard error	.12	.10	.12
	D. Market Portfolio		
Model	−.35	−.33	−.33
Data	−.50	−.30	−.09
Standard error	.11	.13	.13
	E. Value Factor		
Model	−.34	−.45	−.49
Data	−.34	−.44	−.52
Standard error	.09	.09	.12

NOTE.—This table reports the correlation between differences in the rate of innovation (ν_t/M_t) and returns to the market portfolio, the value factor (the return spread between the top and the bottom decile portfolios of stocks sorted on book-to-market ratio), aggregate consumption and investment growth (NIPA), and the consumption of shareholders using the data from Malloy, Moskowitz, and Vissing-Jorgensen (2009). For portfolio returns, we compute the cumulative portfolio return between t and $t + h$. For all other variables, including innovation, we use the log difference between t and $t + h$. We compute correlations at horizons of 1–3 years. Standard errors are computed using Newey-West, with maximum lag length equal to 3 plus the number of overlapping observations. Data period is 1933–2008.

weakly negatively correlated with the consumption growth of shareholders and is essentially uncorrelated with the aggregate consumption growth. Importantly, the magnitudes are comparable, suggesting that the model mechanism does not rely on counterfactually large responses in aggregate consumption growth. Panel C shows that both in the data and in the model, changes in the rate of innovation ν_t/M_t are positively correlated with investment growth. In the data, the estimated correlations are lower than in the model (27%–30% vs. 60%–66%), and the difference is statistically significant. Although there are clearly other drivers of investment growth in the data than technological innovation, the empirical results suggest that the contribution of innovation to investment growth is not trivial.

Last, panels D and E of table 7 compare the correlation between changes in the rate of innovation and asset returns, specifically the market portfolio and the value factor. In the data, changes in innovation are negatively correlated with returns to the market portfolio, although the relation is stronger at the 1-year versus the 3-year frequency. In the model, the relation is also negative and mostly comparable to the data. We also see that the value minus growth factor is negatively correlated with changes in the rate of innovation at all frequencies. Importantly, the magnitude of this correlation is comparable between the data and the model.

In sum, we see that the correlations between the aggregate rate of innovation and key quantities in the model are largely consistent with their empirical counterparts. We view the degree the model can quantitatively replicate these new facts—especially since they are not part of our estimation targets—as providing further support for the main mechanism.

G. *Firm Innovation and Tobin's Q*

We next focus our attention at the more micro level and, specifically, on firm-level outcomes. The model implies that part of the reason why the value premium arises is that growth firms innovate more than value firms. We examine such prediction in this section; our analysis should be viewed as a direct test of equation (30). Specifically, we are interested in whether a firm's valuation ratio (Tobin's Q) predicts the firm's innovative output. We therefore estimate

$$\frac{\nu_{f,t+1}}{M_{f,t+1}} = a + b \log Q_{f,t} + cZ_{f,t} + \varepsilon_{f,t}. \tag{37}$$

Table 8 presents the results. We see that in both the data as well as the model, Tobin's Q is a strong predictor of whether the firm is likely to innovate in the future once we control for its current size. Importantly, the estimated elasticities are comparable between the model and the data. Furthermore, the positive relation survives controlling for whether

TABLE 8
VALUATION RATIOS AND FUTURE INNOVATION: MODEL VERSUS DATA

$\nu_{f,t+1} / M_{f,t+1}$	(1)	(2)	(3)	(4)
log $Q_{f,t}$:				
Model	−.038	.244	.244	.125
Data	−.007	.211***	.169***	.080***
Standard error	.018	.023	.022	.015
log $K_{f,t}$:				
Model		.160	.160	.045
Data		.189***	.243***	.096***
Standard error		.014	.018	.010
$\nu_{f,t+1} / M_{f,t+1}$:				
Model				.569
Data				.657***
Standard error				.062
Observations	70,987	70,987	70,986	65,823
R^2	.055	.194	.309	.571
Fixed effects	Time	Time	Time, industry	Time, industry

NOTE.—This table reports the results of estimating equation (37). We include controls in $Z_{f,t}$ for the firm's current size, time dummies, and lagged values of the dependent variable. We also consider specifications with industry fixed effects, defined at the three-digit Standard Industrial Classification level. The first row reports results in simulated data based on a long sample of 10,000 years. The next two rows report results in Compustat data along with standard errors clustered by firm. In addition to Q, we include controls for firm size ($K_{f,t}$ in the model, Compustat:ppegt in the data) and lagged values of firm innovation. Since firms that never patent may not be a valid control group, we restrict the sample to include only firms that have filed at least one patent.

*** Statistically significant at the 1% level.

the firm successfully innovated in the current period. Since the likelihood of successful innovation is somewhat random, Tobin's Q carries additional information in the model about the firm's current state relative to whether the firm innovated in the past. The same is true in the data.

H. Technological Innovation and Firm Displacement

Having established that growth firms are indeed more likely to innovate than value firms, we next turn to firm profitability. As we see in equations (31) and (32) and figure 5, the model has specific predictions about the response of firm profitability to changes in the technology frontier as a function of the firm's current state (value or growth). Here, we examine these predictions directly. To get sharper estimates of the impact of technological progress on firm profitability, we take advantage of the substantial heterogeneity in innovation outcomes across industries (see, e.g., the working paper version of Kogan et al. 2017). Specifically, we construct a direct analog of ω at the industry level as $\nu_{I \setminus f,t}/M_{I \setminus f,t}$; we aggregate over all firms in industry I, excluding firm f.

We estimate the impact of technology on log firm profitability using the following specification:

$$\log\Pi_{f,t+T} - \log\Pi_{f,t} = \left(a_0 + a_1 G_{f,t}\right)\frac{\nu_{I\setminus f,t}}{M_{I\setminus f,t}} + a_2 G_{f,t} + cZ_{f,t} + \varepsilon_{t+T}. \quad (38)$$

The dependent variable is the change in the firm's log profits between time t and $t + T$. We are interested in the impact of industry innovation $\nu_{I\setminus f,t}/M_{I\setminus f,t}$ on firms with different levels of Tobin's Q. In particular, we classify firms as either value ($G_{f,t} = 0$) or growth ($G_{f,t} = 1$) depending on whether their Tobin's Q falls below or above the industry median at time t.

Table 9 compares estimates of equation (38) across horizons of 1–6 years in the model and in the data. In the data, technological innovation by other firms in the same industry primarily hurts value (low Q) firms. In relative terms, growth firms benefit. These patterns are consistent with the model. In particular, the estimated coefficient a_0 in the data is negative and statistically different from zero, implying that the impact of technological progress on the expected profitability of low-Q firms is negative. Further, the magnitudes are comparable between the model and the data. More importantly, consistent with the model's main mechanism, the estimated coefficient a_1 is positive and statistically significant across horizons, implying that the profits of value firms are more negatively exposed to industry innovation shocks than the profits of growth firms. The magnitudes are somewhat comparable at shorter horizons (1–3 years), but they differ statistically at longer horizons. That is, the model implies somewhat larger heterogeneity between value and growth firms

TABLE 9
TECHNOLOGY SHOCKS AND FIRM PROFITABILITY

	HORIZON				
$\Delta\pi_{f,t\to t+h}$	(1)	(2)	(3)	(4)	(5)
$\nu_{I\setminus f,t} / M_{I\setminus f,t}$:					
Model	−.029	−.044	−.055	−.063	−.069
Data	−.037***	−.036***	−.043***	−.040***	−.043***
Standard error	.004	.005	.007	.008	.009
$(\nu_{I\setminus f,t} / M_{I\setminus f,t}) \times G_{fi}$:					
Model	.012	.032	.049	.065	.078
Data	.024***	.024***	.030***	.032***	.031***
Standard error	.004	.006	.008	.008	.009
Observations	68,818	63,618	59,278	55,321	51,551
R^2	.118	.137	.150	.149	.157

NOTE.—This table summarizes the estimated coefficients a_0 and a_1 from equation (38) in historical data and in simulated data from the model. The vector of controls $Z_{f,t}$ includes the firm's own innovation outcome, the analog of the $dN_{f,t}$ term in equation (32), measured as $\nu_{f,t}/M_{f,t}$; controls for lagged log profits $\pi_{f,t}$ and log size $K_{f,t}$ to be consistent with eqq. (31)–(32); and time and industry dummies. We omit the time and industry dummies in simulated data. We report standard errors clustered by firm. The model estimates are from a long panel (5,000 years) comprised of 1,000 firms. We scale variables to standard deviation units to facilitate comparison between the model and the data.

*** Statistically significant at the 1% level.

at longer horizons than what we see in the data. This difference may be partly driven by the fact that the smallest growth firms in the economy are not in the Compustat database, along with the fact that the less successful firms may exit the sample—the model has no firm exit.

I. Technological Innovation and Inequality

A key feature of our model is that a small subset of households—those that successfully innovate—capture a substantial fraction of the rents from innovation. By contrast, the cost of innovation—the displacement of the installed capital stock—is shared by all market participants. As a result, our model has predictions about the relation between technological innovation and changes in inequality, especially at the very top. Our goal in this section is to examine this implication of the model in the data using empirical measures of inequality.

Our first measure of inequality is based on the difference between average and median consumption in the economy. Recalling figure 4, which shows that improvements in technology have very different effects on the average versus the median future consumption path of an individual household, a natural measure of inequality is the ratio of the average (National Income and Product Accounts [NIPA]) consumption to the median in the cross section of stock-owning households in the CEX. Our second measure of inequality aims to capture the idea that the benefits of innovation are shared by a handful of households. We use the data of Piketty and Saez (2003) to estimate the ratio of the 0.1% top income share to the top 1%. This fractal measure of inequality captures the share of the top 1% that accrues to the top 0.1% and displays similar behavior as the top income shares (Gabaix et al. 2016). To examine the quantitative implications of the model, we construct each of these two measures in simulated data from the model.

For each of the three inequality measures, we estimate the following specification:

$$\begin{aligned}\log\mathrm{INEQ}_T - \log\mathrm{INEQ}_t = a_0 + b_T\left(\log\left(\frac{\nu_T}{M_T}\right) - \log\left(\frac{\nu_t}{M_t}\right)\right) \\ + c_1\mathrm{INEQ}_t + c_2\log\left(\frac{\nu_t}{M_t}\right) + \varepsilon_t,\end{aligned} \tag{39}$$

where ν/M is equal to the average value of new projects (patents) over the total market capitalization. To evaluate the magnitude of the empirical estimates, we also estimate equation (39) in a long simulation of the model (10,000 years).

Figure 7 plots the estimated coefficients b_T across horizons in the data (solid line) and in the model (dotted line). In panel A, we see that changes

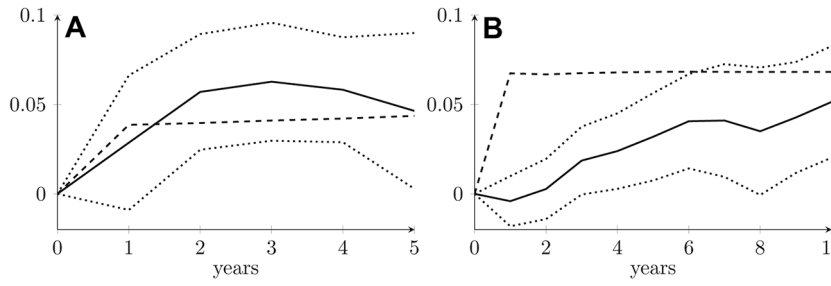


FIG. 7.—Technology shocks and inequality. We plot the estimated coefficient b_1 from equation (39). *A*, Consumption inequality. *B*, Income inequality. For consumption inequality, we compute the ratio of the log aggregate per capita NIPA personal consumption expenditures to the median consumption in the CEX. To be consistent with the model, we estimate the median using only the set of households that are stockholders. Given the short time dimension in the CEX, we focus on horizons T of up to 5 years. For income inequality, we use the ratio of top fractile income shares (0.1% and 1% shares) of Piketty and Saez (2003) that include capital gains. Income in the model is defined as the sum of wage income, capital income (payout), plus payments to innovators. We examine horizons T of 1–10 years. We plot empirical point estimates along with 90% confidence intervals using Newey-West errors. Model results are computed using a long panel of 10 million households over 10,000 years.

in the value of new innovations ν_t/M_t are associated with increases in consumption inequality in the data. Further, the estimated coefficients in the data are quantitatively close to those implied by the model. One concern with using CEX data is that it tends to undersample rich households. Even though using the full consumption distribution of the CEX is problematic for our purposes, the effect of undersampling on the median should be minor. However, we emphasize that, since the full extent of the sample selection biases in the CEX data is unknown, these results should be interpreted with caution.

In panel B, we examine the response of top fractile income shares. The estimated coefficients b_T are mostly positive and statistically significant over horizons of more than 3 years; an increase of 1 standard deviation in ω is associated with an increase in income inequality of about 5% over 10 years. In the model simulations, the increase in income inequality is quantitatively similar in the long run (approximately 6%), but it occurs much faster than in the data. In the model, increases in ω raise the income that accrues to innovators; since the measure of innovating households is very small, increases in ω increase income inequality at the very top of the distribution. However, unlike reality, this process of reallocation occurs instantaneously in the model, and hence the immediate increase in income inequality. That said, existing models have typically the opposite problem; that is, they generate an increase in income inequality that is too slow relative to the data (Gabaix et al. 2016). Hence, we view our model as also contributing to the literature focusing on the rapid increase in income inequality over the last few decades.

VI. Conclusion

We develop a general equilibrium model to study the effects of innovation on asset returns. The main feature of our model is that the benefits from technological progress are not shared symmetrically across all agents in the economy. Specifically, technological improvements partly benefit agents that are key in the creation and implementation of new ideas. As a result, technology shocks also lead to substantial reallocation of wealth among households. Embodied shocks have a large reallocative effect, whereas disembodied shocks have mostly a level effect on household consumption. In equilibrium, shareholders invest in growth firms despite their low average returns, as they provide insurance against increases in the probability of future wealth reallocation. Our model delivers rich cross-sectional implications about the effect of innovation on firms and households that are supported by the data.

Our work suggests several promising avenues for future research. First, labor income in our model is homogeneous, and therefore workers benefit from both types of technological progress. In practice, however, technological advances are often complementary to only a subset of workers' skills. Recent evidence, for example, shows that the job market has become increasingly polarized (Autor, Katz, and Kearney 2006). Thus, quantifying the role of technological progress as a determinant of the risk of human capital may be particularly important. Second, technological progress tends to disrupt traditional methods of production, leading to periods of increased uncertainty. If some agents have preferences for robust control, higher levels of uncertainty will likely increase the agents' demand for insurance against improvements in technology embodied in new vintages. Last, our model implies that claims on any factor of production that can be used across different technology vintages (as, for instance, land) can have an insurance role similar to the one played by growth firms in our current framework. Our model would therefore imply that claims on such factors should have lower equilibrium expected returns.

References

Abel, Andrew B. 1990. "Asset Prices under Habit Formation and Catching Up with the Joneses." *A.E.R.* 80 (2): 38–42.

———. 1999. "Risk Premia and Term Premia in General Equilibrium." *J. Monetary Econ.* 43 (1): 3–33.

Ai, Hengjie, Mariano Massimiliano Croce, and Kai Li. 2013. "Toward a Quantitative General Equilibrium Asset Pricing Model with Intangible Capital." *Rev. Financial Studies* 26 (2): 491–530.

Asquith, Paul, and E. Han Kim. 1982. "The Impact of Merger Bids on the Participating Firms' Security Holders." *J. Finance* 37 (5): 1209–28.

Atkeson, Andrew, and Patrick J. Kehoe. 2005. "Modeling and Measuring Organization Capital." *J.P.E.* 113 (5): 1026–53.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *A.E.R.* 96 (2): 189–94.
Bansal, Ravi, Dana Kiku, and Amir Yaron. 2012. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices." *Critical Finance Rev.* 1 (1): 183–221.
Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *J. Finance* 59 (4): 1481–509.
Beeler, Jason, and John Y. Campbell. 2012. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *Critical Finance Rev.* 1 (1): 141–82.
Berndt, Ernst R. 1990. "Energy Use, Technical Progress and Productivity Growth: A Survey of Economic Issues." *J. Productivity Analysis* 2 (1): 67–83.
Bhide, A. V. 1999. *The Origin and Evolution of New Businesses*. Oxford Univ. Press, Oxford.
Blanchard, Olivier J. 1985. "Debt, Deficits, and Finite Horizons." *J.P.E.* 93 (2): 223–47.
Bolton, Patrick, Neng Wang, and Jinqiang Yang. 2015. "A Theory of Liquidity and Risk Management Based on the Inalienability of Risky Human Capital." Working paper, Columbia Univ.
Campbell, John Y., and John Cochrane. 1999. "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *J.P.E.* 107 (2): 205–51.
Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *A.E.R.* 102 (6): 2981–3003.
Clark, Andrew E., and Andrew J. Oswald. 1996. "Satisfaction and Comparison Income." *J. Public Econ.* 61 (3): 359–81.
Clementi, Gian Luca, and Berardino Palazzo. 2019. "Investment and the Cross-Section of Equity Returns." *J. Finance* 74 (1): 281–321.
Cochrane, J. H. 2005. *Financial Markets and the Real Economy. Foundations and Trends in Finance* 1 (1): 1–101.
Constantinides, George M. 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle." *J.P.E.* 98 (3): 519–43.
Constantinides, George M., and Darrell Duffie. 1996. "Asset Pricing with Heterogeneous Consumers." *J.P.E.* 104 (2): 219–40.
DeMarzo, Peter M., Ron Kaniel, and Ilan Kremer. 2008. "Relative Wealth Concerns and Financial Bubbles." *Rev. Financial Studies* 21 (1): 19–50.
Dew-Becker, Ian. 2017. "How Risky Is Consumption in the Long-Run? Benchmark Estimates from a Robust Estimator." *Rev. Financial Studies* 30 (2): 631–66.
Duesenberry, J. S. 1949. *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard Univ. Press.
Duffie, Darrell, and Larry G. Epstein. 1992. "Stochastic Differential Utility." *Econometrica* 60 (2): 353–94.
Easterlin, Richard A. 1974. "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." *Nations and Households Econ. Growth* 89:89–125.
Eisfeldt, Andrea L., and Dimitris Papanikolaou. 2013. "Organization Capital and the Cross-Section of Expected Returns." *J. Finance* 68 (4): 1365–1406.
Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4): 937–69.
Fama, Eugene F., and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." *J. Finance* 47 (2): 427–65.
———. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *J. Financial Econ.* 33 (1): 3–56.

Fisher, Jonas D. M. 2006. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks." *J.P.E.* 114 (3): 413–51.
Frydman, Cary. 2015. "What Drives Peer Effects in Financial Decision-Making? Neural and Behavioral Evidence." Working paper, Univ. Southern California.
Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The Dynamics of Inequality." *Econometrica* 84 (6): 2071–111.
Garleanu, Nicolae, Leonid Kogan, and Stavros Panageas. 2012. "Displacement Risk and Asset Returns." *J. Financial Econ.* 105 (3): 491–510.
Garleanu, Nicolae, Stavros Panageas, and Jianfeng Yu. 2012. "Technological Growth and Asset Pricing." *J. Finance* 67 (4): 1265–92.
Gentzkow, Matthew, and Jesse M. Shapiro. 2017. "Measuring the Sensitivity of Parameter Estimates to Estimation Moments." *Q.J.E.* 132 (4): 1553–92.
Gomes, Joao F., Leonid Kogan, and Lu Zhang. 2003. "Equilibrium Cross Section of Returns." *J.P.E.* 111 (4): 693–732.
Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." *A.E.R.* 87 (3): 342–62.
Hart, Oliver, and John Moore. 1994. "A Theory of Debt Based on the Inalienability of Human Capital." *Q.J.E.* 109 (4): 841–79.
Ingram, Beth Fisher, and Bong-Soo Lee. 1991. "Simulation Estimation of Time-Series Models." *J. Econometrics* 47 (2–3): 197–205.
Jensen, J. Bradford, Robert H. McGuckin, and Kevin J. Stiroh. 2001. "The Impact of Vintage and Survival on Productivity: Evidence from Cohorts of U.S. Manufacturing Plants." *Rev. Econ. and Statis.* 83 (2): 323–32.
Jermann, Urban J. 1998. "Asset Pricing in Production Economies." *J. Monetary Econ.* 41 (2): 257–75.
Kaltenbrunner, Georg, and Lars A. Lochstoer. 2010. "Long-Run Risk through Consumption Smoothing." *Rev. Financial Studies* 23 (8): 3190–224.
Kogan, Leonid, and Dimitris Papanikolaou. 2013. "Firm Characteristics and Stock Returns: The Role of Investment-Specific Shocks." *Rev. Financial Studies* 26 (11): 2718–59.
———. 2014. "Growth Opportunities, Technology Shocks, and Asset Prices." *J. Finance* 69 (2): 675–718.
Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *Q.J.E.* 132 (2): 665–712.
Korteweg, Arthur, and Stefan Nagel. 2016. "Risk-Adjusting the Returns to Venture Capital." *J. Finance* 71 (3): 1437–70.
Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh. 2011. "Technological Change and the Growing Inequality in Managerial Compensation." *J. Financial Econ.* 99 (3): 601–27.
Luttmer, Erzo F. P. 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." *Q.J.E.* 120 (3): 963–1002.
Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jorgensen. 2009. "Long-Run Stockholder Consumption Risk and Asset Returns." *J. Finance* 64 (6): 2427–79.
Nagel, Stefan, Serhiy Kozak, and Shrihari Santosh. 2014. "Interpreting Factor Models." Working paper, Univ. Michigan.
O'Rand, A. M., and J. C. Henretta. 2000. *Age and Inequality: Diverse Pathways through Later Life*. Boulder, CO: Westview.
Papanikolaou, Dimitris. 2011. "Investment Shocks and Asset Prices." *J.P.E.* 119 (4): 639–85.

Pastor, Lubos, and Pietro Veronesi. 2009. "Technological Revolutions and Stock Prices." *A.E.R.* 99 (4): 1451–83.
Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Q.J.E.* 118 (1): 1–41.
Poterba, James M., and Andrew A. Samwick. 1995. "Stock Ownership Patterns, Stock Market Fluctuations, and Consumption." *Brookings Papers on Econ. Activity* 26 (2): 295–372.
Rauh, Joshua D., and Amir Sufi. 2011. "Explaining Corporate Capital Structure: Product Markets, Leases, and Asset Similarity." *Rev. Finance* 16 (1): 115–55.
Rayo, Luis, and Gary S. Becker. 2007. "Evolutionary Efficiency and Happiness." *J.P.E.* 115:302–37.
Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein. 1985. "Persuasive Evidence of Market Inefficiency." *J. Portfolio Management* 11 (3): 9–16.
Roussanov, Nikolai. 2010. "Diversification and Its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status." *J. Finance* 65 (5): 1755–88.
Rouwenhorst, Geert K. 1995. "Asset Pricing Implications of Equilibrium Business Cycle Models." In *Frontiers of Business Cycle Research*, edited by Thomas F. Cooley, chap. 10. Princeton, NJ: Princeton Univ. Press.
Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Q.J.E.* 131 (2): 519–78.
Solow, Robert M. 1960. "Investment and Technical Progress." In *Mathematical Methods in the Social Sciences*, edited by Kenneth J. Arrow, Amuel Karlin, and Patrick Suppes, 89–104. Stanford, CA: Stanford Univ. Press.
Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2018. "Firming Up Inequality." *Q.J.E.* 134 (1): 1–50.
Titman, Sheridan, K. C. John Wei, and Feixue Xie. 2004. "Capital Investments and Stock Returns." *J. Financial and Quantitative Analysis* 39 (4): 677–700.
van Binsbergen, Jules, Michael Brandt, and Ralph Koijen. 2012. "On the Timing and Pricing of Dividends." *A.E.R.* 102 (4): 1596–618.
Wachter, Jessica A. 2013. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *J. Finance* 68 (3): 987–1035.