Displacement Risk and Asset Returns: Extended Appendix

Contents

1 Labor Input as a Composite Service and Human-Capital Depreciation

This section provides a more elaborate model of the labor market, which reproduces the path of labor income over an agent's life, as postulated in equations (10) and (11). The main difference between the baseline model and the model of this section is that the labor income process results from general-equilibrium wage effects, rather than assumptions on agents' endowments of labor efficiency units.

To draw this distinction, we assume that workers' efficiency units are only affected by aging and experience. Specifically, workers endowments of labor efficiency units evolve deterministically over their life according to $h_{t,s} = \overline{h}\,(1+\delta)^{t-s}$. Thus, the innovation shocks u_t no longer have any effect on agents' endowment of labor efficiency units.

Assume, moreover, that labor is not a homogenous service. Instead, the units of labor that enter the production function of final goods and intermediate goods are measured in terms of a composite service, which is a constant-elasticity-of-substitution (CES) aggregator of the labor efficiency units provided by workers belonging to different cohorts. Specifically, one unit of (composite) labor is given by

$$L_t = \left(\sum_{s=-\infty}^{t} v_{t,s}^{\frac{1}{b}} \left(l_{t,s} \right)^{\frac{b-1}{b}} \right)^{\frac{b}{b-1}}, \tag{1.1}$$

where $l_{t,s}$ denotes the labor input of cohort s at time t, $v_{t,s} > 0$ controls the relative importance of this input and $b > 0$ is the elasticity of substitution. The production function of final goods continues to be given by (3) and it still takes one unit of the composite labor service to produce one unit of the intermediate good. Equation (1.1) captures the idea that different cohorts have different skills and hence they are imperfect substitutes in the production process. Next, we let

$$v_{t,s}^{\frac{1}{b}} \equiv (1-\phi)^{\left(\frac{b-1}{b}\right)} q_{t,s} h_{t,s}^{\frac{1}{b}}. \tag{1.2}$$

Before proceeding, we note that using (1.2) inside (1.1), recognizing that in equilibrium $l_{t,s} =$

$h_{t,s}$, and noting that $\sum_{s=-\infty}^{t} q_{t,s}h_{t,s} = 1$ implies that the aggregate supply of (composite) labor efficiency units is constant and equal to $(1-\phi)$.

Since labor inputs by agents belonging to different cohorts are imperfect substitutes, we need to solve for the equilibrium wage $w_{t,s}$ of each separate cohort. This process is greatly facilitated by first constructing a "wage index", i.e., taking a set of cohort-specific wages as given, and then minimizing (over cohort labor inputs) the cost of obtaining a single unit of the composite labor input. As is well established in the literature, this wage index for CES production functions is given by

$$\overline{w}_t = \left(\sum_{s=-\infty}^{t} v_{t,s} \left(w_{t,s} \right)^{1-b} \right)^{\frac{1}{1-b}} .$$

With this wage index at hand, the cohort-specific input demands for a firm demanding a total of L_t units of the composite good are given by

$$w_{t,s} = \overline{w}_t v_{t,s}^{\frac{1}{b}} \left(\frac{l_{t,s}}{L_t} \right)^{-\frac{1}{b}} . \tag{1.3}$$

It is now straightforward to verify that an equilibrium in such an extended model can be determined by setting $\overline{w}_t = w_t$ (where w_t is given by [46]) and then obtaining the cohort-specific wages by setting $l_{t,s} = h_{t,s}$, and $L_t = (1-\phi)$ in equation (1.3) and solving for $w_{t,s}$. To see this, note that by making these substitutions and using (1.2) inside (1.3) leads to the per-worker income process

$$\frac{w_{t,s}h_{t,s}}{(1-\phi)} = \overline{w}_t q_{t,s} , \tag{1.4}$$

which coincides with the labor income process in the baseline model. Furthermore, by setting $w_t = \overline{w}_t$, the market for total (composite) labor units clears by construction, whereas the cohort specific wages implied by (1.4) clear all cohort specific labor markets, since they satisfy equation (1.3) for all markets.

2 A Multi-Sector Extension

It is straightforward to extend the model to allow for multiple sectors, with potentially different degrees of innovation within each sector. Such an extension can help illustrate that even when technological progress is different accross industries, the value premium is likely to be particularly salient *within* industries, as it is in the data.

To introduce such an extension, we modify the baseline setup, so that the production of the final good is given by

$$Y_t = Z_t \left(L_t^F\right)^{1-(\alpha_1+\alpha_2)} \left(\int_0^{A_t} x_{j,t}^{\alpha_1} dj\right) \left(\int_0^{B_t} \widetilde{x}_{j,t}^{\alpha_2} dj\right), \tag{2.1}$$

where $\alpha_1 > 0, \alpha_2 > 0$, and $\alpha_1 + \alpha_2 < 1$, and $x_{j,t}$ denotes the intermediate input j in sector A and $\widetilde{x}_{j,t}$ denotes the intermediate input j in sector B. (To simplify the exposition and avoid inessential notation, specification (2.1) implicitly sets the weights $\omega_{j,t}$ on the intermediate goods equal to one). A_t and Z_t evolve as in the baseline version of the model and B_t evolves similarly to A_t, i.e.,

$$\log B_{t+1} = \log B_t + \widetilde{u}_{t+1},$$

where $\widetilde{u}_{t+1}$ is a non-negative random variable that captures technological advancements in sector B. We allow the shocks u_t and $\widetilde{u}_t$ to be correlated. At each point in time t, the representative final-good firm chooses L_t^F, $x_{j,t}$, and $\widetilde{x}_{j,t}$ so as to maximize its profits

$$\pi_t^F = \max_{L_t^F, x_{j,t}, \widetilde{x}_{j,t}} \left\{ Y_t - \int_0^{A_t} p_{j,t} x_{j,t} dj - \int_0^{B_t} \widetilde{p}_{j,t} \widetilde{x}_{j,t} dj - w_t L_t^F \right\}, \tag{2.2}$$

where $p_{j,t}$ and $\widetilde{p}_{j,t}$ are the prices of intermediate goods in sectors A and B, respectively, and w_t is the prevailing wage (per efficiency unit of labor).

Production of intermediate goods (in either sector) still takes the simple form described in the paper (i.e., it takes one unit of labor to produce one unit of intermediate good j, irrespective of the sector). Accordingly, the total labor demand of both sectors is

$$L_t^I = \int_0^{A_t} x_{j,t} dj + \int_0^{B_t} \widetilde{x}_{j,t} dj. \tag{2.3}$$

Finally, to simplify matters, we assume that new firms are specific to sectors and can own only sector-A or sector-B blueprints (but not both). Differentiating (2.2) leads to the following two first-order conditions with respect to $x_{j,t}$ and $\widetilde{x}_{j,t}$.

$$
\begin{aligned}
x_{j,t} &= \left[\frac{p_{j,t}}{\alpha_1 Z_t \left(L_t^F\right)^{1-(\alpha_1+\alpha_2)} \left(\int_0^{B_t} \widetilde{x}_{j,t}^{\alpha_2} dj\right)}\right]^{\frac{1}{\alpha_1-1}}, && (2.4)\\
\widetilde{x}_{j,t} &= \left[\frac{\widetilde{p}_{j,t}}{\alpha_2 Z_t \left(L_t^F\right)^{1-(\alpha_1+\alpha_2)} \left(\int_0^{A_t} x_{j,t}^{\alpha_1} dj\right)}\right]^{\frac{1}{\alpha_2-1}}. && (2.5)
\end{aligned}
$$

Maximizing the profits of intermediate-good firms leads to the same first-order condition as in the baseline version of the model, namely:

$$
\begin{aligned}
p_{j,t} &= \frac{w_t}{\alpha_1}, && (2.6)\\
\widetilde{p}_{j,t} &= \frac{w_t}{\alpha_2}. && (2.7)
\end{aligned}
$$

Combining (2.4) with (2.6) and (2.5) with (2.7) and using the definition of Y_t leads to

$$
\begin{aligned}
x_{j,t} &= \left[\frac{w_t}{\alpha_1^2 Y_t}\left(\int_0^{A_t} x_{j,t}^{\alpha_1} dj\right)\right]^{\frac{1}{\alpha_1-1}} && (2.8)\\
\widetilde{x}_{j,t} &= \left[\frac{w_t}{\alpha_2^2 Y_t}\left(\int_0^{B_t} \widetilde{x}_{j,t}^{\alpha_2} dj\right)\right]^{\frac{1}{\alpha_2-1}}. && (2.9)
\end{aligned}
$$

Since all intermediate goods within a sector face the same demand curve and the same cost structure, we look for a symmetric equilibrium, in which $x_{j,t} = x_t$ and $\widetilde{x}_{j,t} = \widetilde{x}_t$. Under this supposition, $\int_0^{A_t} x_{j,t}^{\alpha_1} dj = A_t x_t^{\alpha_1}$ and $\int_0^{B_t} \widetilde{x}_{j,t}^{\alpha_2} dj = B_t \widetilde{x}_t^{\alpha_2}$, so that equations (2.8) and (2.9) simplify to

$$
\begin{aligned}
x_t &= \alpha_1^2 \left(\frac{Y_t}{w_t}\right)\frac{1}{A_t}, && (2.10)\\
\widetilde{x}_t &= \alpha_2^2 \left(\frac{Y_t}{w_t}\right)\frac{1}{B_t}. && (2.11)
\end{aligned}
$$

Finally, the final-good firm's first-order condition with respect to labor gives $(1-\alpha_1-\alpha_2) Y_t = w_t L_t^F$, which implies that

$$
\frac{Y_t}{w_t} = \frac{L_t^F}{1-\alpha_1-\alpha_2}. \tag{2.12}
$$

Labor-market clearing can be expressed as

$$
L_t^F + A_t x_t + B_t \widetilde{x}_t = (1-\phi). \tag{2.13}
$$

Using (2.12) inside (2.10) and (2.11), and then using the resulting expressions inside (2.13) and solving for L_t^F leads to

$$L_t^F = \frac{1-\alpha_1-\alpha_2}{1-\alpha_1-\alpha_2+\alpha_1^2+\alpha_2^2}\left(1-\phi\right). \tag{2.14}$$

Combining (2.10), (2.11), (2.12), and (2.14) gives

$$x_t = \frac{\alpha_1^2}{1-\alpha_1-\alpha_2+\alpha_1^2+\alpha_2^2}\frac{1-\phi}{A_t}, \tag{2.15}$$

$$\widetilde{x}_t = \frac{\alpha_2^2}{1-\alpha_1-\alpha_2+\alpha_1^2+\alpha_2^2}\frac{1-\phi}{B_t}. \tag{2.16}$$

Combining (2.14) with (2.15) and (2.16) yields

$$Y_t = \frac{\left(1-\phi\right)\left(1-\alpha_1-\alpha_2\right)^{1-(\alpha_1+\alpha_2)}\alpha_1^{2\alpha_1}\alpha_2^{2\alpha_2}}{1-\alpha_1-\alpha_2+\alpha_1^2+\alpha_2^2}Z_t A_t^{1-\alpha_1}B_t^{1-\alpha_2}. \tag{2.17}$$

Equation (2.17) states that output is proportional to $Z_t A_t^{1-\alpha_1}B_t^{1-\alpha_2}$. From a practical perspective, this implies that the model with multiple sectors behaves like a single-sector model, where the technology shock u_t is replaced by a weighted sum of the technology shocks to the two sectors.[1] In particular all the conclusions regarding the displacement effect are unaltered, with the understanding that the shock u_t in the baseline model is now an appropriate weighted average of the shocks in the two sectors.

Even though the extension to multiple sectors adds little in terms of the model's general-equilibrium properties, it helps clarify that even when technological progress is concentrated in one sector, most of the cross-sectional differences in returns manifest themselves within a sector, rather than across sectors. To see this, note first that per-firm profits in sector A

[1] To see this, note that

$$\Delta \log Y_t = \varepsilon_t + \left(1-\alpha_1\right)u_t + \left(1-\alpha_2\right)\widetilde{u}_t.$$

Defining $u_t^* = \left(1-\alpha_1\right)u_t + \left(1-\alpha_2\right)\widetilde{u}_t$ shows that the output growth in the multisector model is identical to the one in the single-sector model, with u_t^* defined appropriately to capture the total effect of all displacement shocks.

and B are given by

$$\pi_t^A = \alpha_1 (1 - \alpha_1) \left(\frac{Y_t}{A_t} \right),$$
$$\pi_t^B = \alpha_2 (1 - \alpha_2) \left(\frac{Y_t}{B_t} \right).$$

Now suppose that technological advancements are concentrated in one sector (say, sector A), so that u_t is random, but $\widetilde{u}_t$ is a constant. Take any stock in sector B. Since $\widetilde{u}_t$ is non-random, there will be no difference between the rates of return of different firms in sector B.[2] By contrast, stocks in sector A exhibit a non-trivial value premium, with "pure" growth options in sector A having a lower expected return than sector-B stocks (since they act as a hedge against the u-shock) and pure value stocks in sector A having higher expected returns than sector-B stocks. As a result, the "HML" factor in this economy is driven exclusively by return differentials *within* sector A.

3 Robustness of the Cohort Methodology: Implications for the Permanent Component of the Stochastic Discount Factor

In this section we show two results: First, the cohort methodology that we use relies only on a minimal set of assumptions, which are shared by several overlapping generations models. Second, even when these identifying assumptions are relaxed, our main conclusions remain

[2]To see this, let R_t^a denote the return of a "pure" asset in place in sector B, and R_t^o the return on a "pure" growth option in sector B. The definitions of R_t^a and R_t^o in the paper imply that $\log R_{t+1}^a - \log R_{t+1}^o$ is a non-random constant. Indeed, it is zero, since

$$1 = E_t \left(e^{\log R_{t+1}^a} \frac{\xi_{t+1}}{\xi_t} \right) = E_t \left(e^{\log R_{t+1}^a - \log R_{t+1}^o} \times e^{\log R_{t+1}^o} \frac{\xi_{t+1}}{\xi_t} \right)$$
$$= e^{\log R_{t+1}^a - \log R_{t+1}^o} \times E_t \left(e^{\log R_{t+1}^o} \frac{\xi_{t+1}}{\xi_t} \right) = e^{\log R_{t+1}^a - \log R_{t+1}^o}$$

and hence $\log R_{t+1}^a = \log R_{t+1}^o$.

unaltered, provided that we interpret our results as pertaining to the permanent component of the stochastic discount factor.

Specifically, to see the first point, take any OLG model satisfying the following two assumptions:

1. Existing agents' consumption satisfy Euler equations of the form

$$\beta^{t-s}\left(\frac{c_{t,s}}{c_{s,s}}\right)^{-\gamma} = \frac{\xi_t}{\xi_s}, \tag{3.1}$$

   where ξ_t is the stochastic discount factor, and the rest of the notation is explained in the text. (For the purposes of the extended appendix we ignore, for simplicity, the type of external habit formation that is allowed in the paper.)

2. An incoming generation's consumption is a stationary fraction of aggregate output, i.e., $z_s \equiv \frac{c_{s,s}}{Y_s}$ is a stationary process.

Just based on the above two assumptions, the permanent component of a time-series of consumption cohort effects can help uncover the *permanent* component of the stochastic discount factor that is due to displacement.

To see why, suppose that we first take logarithms in equation (3.1), and then estimate cohort effects a_s. Mathematically, the identification of cohort effects relies on the difference in the time-t consumption of different cohorts:[3]

$$\begin{aligned} \Delta a_{s+1} &\equiv a_{s+1} - a_s \equiv \log\left(c_{t,s+1}\right) - \log\left(c_{t,s}\right) && (3.2)\\ &= \Delta \log z_{s+1} + \Delta \log Y_{s+1} - \frac{1}{\gamma}\Delta \log \xi_{s+1} + \frac{1}{\gamma}\log\beta, && (3.3) \end{aligned}$$

where the second line follows upon substituting (3.1) into (3.2). Using (3.3) to solve for $\Delta \log \xi_{s+1}$ and computing $\sum_{s=0}^{t-1} \Delta \log \xi_{s+1}$ leads to

$$\log \xi_t = B_0 - \gamma \log z_t - \gamma \log Y_t + \gamma a_t + t \log \beta,$$

[3] If we include also age effects, then we need to compute second differences (with respect to s) of $\log c_{t,s}$ as we explain in detail the paper. For simplicity, we ignore age effects here, but we allow for them in the paper.

where B_0 is a constant given by $\log \xi_0 + \gamma \log z_0 Y_0 - \gamma a_0 - \log \beta$. Using the standard, Beveridge-Nelson definition of the permanent component of a time-series[4], and noting that by assumption 2, z_t is stationary implies that

$$\text{perm. comp.}\,(\log \xi_t) = \underbrace{-\gamma \text{perm. comp.}\,(\log Y_t)}_{\text{“rising-tide component”}} + \underbrace{\gamma \text{perm. comp.}\,(a_t)}_{\text{“intergenerational distribution”}}. \tag{3.4}$$

Equation (3.4) shows that assumptions 1 and 2 directly imply that the cohort effects a_t identify variations in the permanent component of the log-stochastic discount factor that are distinct from variations caused by aggregate output shocks.

We would like to make three observations about the decomposition of the stochastic discount factor derived in (3.4).

First, the above calculations apply to any OLG model, as long as it satisfies the two (rather basic) assumptions outlined above. Second, equation (3.4) does not depend on any theory about what drives the lack of intergenerational risk sharing. This is attractive since it separates the identification of distributional disturbances due to lack of intergenerational risk sharing from the exact “story” driving such imperfect risk sharing across generations. Third, equation (3.4) shows that cohort effects identify distributional shocks that potentially matter for asset pricing, since they are reflected in the stochastic discount factor.

One may argue that equation (3.1) may be common across many models, but fails in the data. Indeed, in the data there is evidence that risk is imperfectly shared even among existing generations. Such imperfect risk sharing between coexisting cohorts is most likely related to borrowing constraints, limited participation, etc., which drive a wedge between $\beta^{t-s}\left(\frac{c_{t,s}}{c_{s,s}}\right)^{-\gamma}$ and $\frac{\xi_t}{\xi_s}$. This wedge is more likely to be prevalent early in life, when an agent has more human capital rather than financial wealth. We proceed now to argue that our previous conclusions on the the relationship between the permanent components of cohort effects and the stochastic discount factor are unaffected by such a wedge, as long as it is

[4]The Beveridge Nelson definition states that for a series y_t the permanent component is defined as

$$\lim_{T\to\infty} \left[E_t y_{t+T} - TE\Delta y_{t+1}\right].$$

transient.

To study the implications of such a wedge for our identification strategy, suppose that the consumption of an agent belonging to cohort s is given by

$$\Delta \log c_{t+1,s} = \begin{cases} \frac{1}{\gamma}\log\beta - \frac{1}{\gamma}\Delta\log\xi_{t+1} - \frac{1}{\gamma}\eta_{t+1} \text{ if } t-s=1 \\ \frac{1}{\gamma}\log\beta - \frac{1}{\gamma}\Delta\log\xi_{t+1} \text{ otherwise,} \end{cases} \tag{3.5}$$

where η_t is an Euler-equation residual. We remain agnostic as to the underlying friction. Equation (3.5) is simply meant to generalize (3.1) by leaving room for deviations from the Euler equation early in life.

Iterating (3.5) forward leads to a generalized version of (3.1):

$$\beta^{t-s}\left(\frac{c_{t,s}}{c_{s,s}}\right)^{-\gamma} = \frac{\xi_t}{\xi_s}e^{\eta_{s+1}}.$$

Now, postulating again that $z_s = \frac{c_{s,s}}{Y_s}$ is stationary and simply repeating the same calculations as the ones we gave above leads to

$$\Delta a_{s+1} = \Delta\log z_{s+1} + \Delta\log Y_{s+1} - \frac{1}{\gamma}\Delta\log\xi_{s+1} + \frac{1}{\gamma}\log\beta + \frac{1}{\gamma}\Delta\eta_{s+2}. \tag{3.6}$$

Even though equation (3.6) is different from (3.2), equation (3.4) still holds. Indeed, computing $\sum_{s=0}^{t-1}\Delta\log\xi_{s+1}$ from (3.6) and using that z_s and η_s are both stationary we have once again

$$\text{perm. comp.}(\log\xi_t) = \underbrace{-\gamma\text{perm. comp.}(\log Y_t)}_{\text{“rising tide component”}} + \underbrace{\gamma\text{perm. comp.}(a_t)}_{\text{“intergenerational distribution”}}, \tag{3.7}$$

which is identical to (3.4). Equation (3.7) shows that our conclusions about the relationship between the permanent component of the stochastic discount factor and the permanent components of consumption cohort effects is not affected by the presence of transient deviations from Euler relationships in an agent's life.

To give some economic substance to what the first-period Euler residual (η_s) may capture, in the next section of this extended appendix we present a concrete example where agents do not participate in asset markets in the first period of their lives. That example illustrates in a specific case that lack of risk sharing between existing generations may affect the "short-run" dynamics of the stochastic discount factor, but does not impact the long-run relationship (3.7).

The same conclusion would hold if we extended the above example to allow for deviations from the Euler equation over the agent's life-time. In that case the first line of (3.5) would not only apply at age $t - s = 1$, but also at $t - s = 2$, 3, etc. To preserve (3.7), the main assumption we would have to make in this case is that the sum of the standard deviations of these residuals is finite over the life cycle. For instance, that would automatically happen if these Euler-equation residuals are present while an agent is young and has substantial human capital, and then "dissipate" later in life as the agent accumulates financial wealth.

To conclude, incomplete risk sharing between existing cohorts will in general affect the dynamics of the stochastic discount factor; however, under reasonable economic assumptions, the permanent component of the stochastic discount factor is still correctly identified by our methodology. Importantly, the permanent component of the stochastic discount factor controls the pricing of risk for long-run returns. Therefore, as long as we interpret our results as pertaining to the properties of long-run returns, our conclusions are robust to the introductions of frictions preventing perfect risk sharing between existing generations.

4 Imperfect Consumption Correlation across Existing Cohorts: an Example

We simplify some aspects of the model for tractability. One of the stylized assumptions is that innovating agents receive their blueprints "at birth." In reality, though, it takes time to start a firm, and each cohort of agents does not innovate simultaneously. Moreover, innovation shocks u_t are more likely to follow a moving-average process rather than being independent, as we assume. We provide a simple example to illustrate why such frictions and perturbations of the baseline model are unlikely to affect our conclusions about the long-run properties of the model-implied asset returns.

Suppose that all agents are born as workers with an initial endowment of labor efficiency units of $\overline{h}\,(1-\phi)\,q_{s,s}$. Furthermore, suppose that a fraction ϕ of them become entrepreneurs in the second period of their lives and permanently drop out of the workforce, whereas the ones that remain workers have an endowment of labor efficiency units equal to the baseline model from the second period of their lives onward, namely $\overline{h}(1+\delta)^{t-s}q_{t,s}$ for all $t \geq s+1$.[5] Finally, assume that agents can only access financial markets in the second period of their lives, while in the first period they consume their wage income. These assumptions capture the idea that an agent's "birth" cohort and the date at which that agent innovates may not coincide. Moreover, exclusion from markets captures in a stylized manner the idea that the agent cannot smooth consumption between the "birth" date and the innovation arrival date — say, due to borrowing constraints.

Repeating the argument of Section 3.2, the equilibrium stochastic discount factor in this modified setup is

$$\frac{\xi_{t+1}}{\xi_t} = \beta \left(\frac{Y_{t+1}}{Y_t}\right)^{-1+\psi(1-\gamma)} \hat{v}(u_{t+1}, u_t)^{-\gamma},$$

where

$$\hat{v}(u_{t+1}, u_t) = (1-\lambda)^{-1}\left(1 - \frac{\lambda y_t}{C_t}\right)^{-1}\left(1 - \lambda\,(1-\lambda)\sum_{i\in\{w,e\}} \phi^i \frac{c^i_{t+1,t}}{C_{t+1}} - \lambda \frac{y_{t+1}}{C_{t+1}}\right)$$

and y_t denotes an agent's initial wage income. Furthermore, the same reasoning as in the proof of Lemma 4 implies that the variance of the permanent component of log consumption cohort effects equals $Var\left(\hat{v}(u_{t+1}, u_t)\right)$.

This simple example illustrates the fact that the frictions affecting agents' life-cycle of earnings change the transitory dynamics of cohort effects, returns, and the stochastic discount factor. Such frictions do not alter our main qualitative conclusion that the permanent component of cohort effects captures the permanent component of the displacement factor, as reflected in the stochastic discount factor.

[5] Note that since agents are born with $\overline{h}\,(1-\phi)\,q_{s,s}$ rather than $\overline{h}q_{s,s}$ efficiency units, the total supply of labor efficiency units remains equal to the baseline model.

5 Robustness of the Cohort-Effect Methodology to Sluggish Adjustment in Consumption

The CEX is a repeated cross section. In the CEX one can observe the consumption of an individual household for about a year, so that the observed consumption growth per household is about three quarters. The set of households that are observed changes from cross-section to cross-section. This poses a problem for asset-pricing estimations, if consumption does not adjust to news over short-run intervals. To frame ideas, we shall assume one specific friction, namely inattention in the spirit of Duffie and Sun (1990), Gabaix and Laibson (2002), and Abel et al. (2010). In these papers agents observe the stock market infrequently, and hence an agent's consumption does not immediately respond to aggregate news.

To study the implications of such a short-run friction in our framework, we study a stylized framework of inattention, based on the key insights of that literature. Specifically, we assume that a fraction q of existing agents observes aggregate outcomes and participates in Arrow-Debreu markets every period of their life, while the remaining agents observe aggregate outcomes and participate in Arrow-Debreu markets only in even periods of their life. We will refer to the first set of agents as "attentive" agents and the latter as "inattentive" agents. All agents participate in the money market every period of their life. To expedite the presentation, we assume that all agents have CRRA preferences.

The consumption of attentive agents satisfies the usual stochastic Euler relationship

$$\frac{c_{t+1,s}^{(A)}}{c_{t,s}^{(A)}} = \left(\beta^{-1}\frac{\xi_{t+1}}{\xi_t}\right)^{-\frac{1}{\gamma}} \tag{5.1}$$

in all periods of their life. (Throughout, the superscript I to refer to inattentive agents and A to attentive agents.) However, the consumption of inattentive agents satisfies a stochastic Euler relationship only between even periods of their life:

$$\frac{c_{t+2,s}^{(I)}}{c_{t,s}^{(I)}} = \left(\beta^{-2}\frac{\xi_{t+2}}{\xi_t}\right)^{-\frac{1}{\gamma}} \text{ if } \frac{t-s}{2} \in \{0, 1, 2...\}. \tag{5.2}$$

When $t-s$ is odd, the consumption of inattentive agents is given by

$$\frac{c_{t,s}^{(I)}}{c_{t-1,s}^{(I)}}=\left(\beta^{-1}E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right)^{-\frac{1}{\gamma}}. \tag{5.3}$$

Note that when $t-s$ is odd, the consumption of inattentive agents satisfies a deterministic Euler equation. In particular, consumption in an odd period of an agent's life depends deterministically on information available at the preceeding even period, consistent with the notion of inattentiveness: The consumer observes her wealth and assets in an even period of life; during those periods he makes a deterministic optimal plan that characterizes her consumption in the ensuing odd period.

An implication of equations (5.1)–(5.3) is that the average marginal rate of substitution of existing agents does not provide a valid pricing kernel:

$$q\beta\left(\frac{c_{t+1,s}^{(A)}}{c_{t,s}^{(A)}}\right)^{-\gamma}+(1-q)\,\beta\left(\frac{c_{t+1,s}^{(I)}}{c_{t,s}^{(I)}}\right)^{-\gamma}\neq\frac{\xi_{t+1}}{\xi_t}.$$

Cohort analysis can still help recover the stochastic discount factor. Iterating equation (5.1) gives

$$\frac{c_{t,s}^{(A)}}{c_{s,s}^{(A)}}=\left(\beta^{-(t-s)}\frac{\xi_t}{\xi_s}\right)^{-\frac{1}{\gamma}},$$

while equations (5.2) and (5.3) lead to

$$\frac{c_{t,s}^{(I)}}{c_{s,s}^{(I)}}=\begin{cases}\left(\beta^{-(t-s)}\frac{\xi_t}{\xi_s}\right)^{-\frac{1}{\gamma}} & \text{if } \frac{t-s}{2}\in\{0,1,2...\}\\ \left(\beta^{-(t-s)}\frac{\xi_{t-1}}{\xi_s}E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right)^{-\frac{1}{\gamma}} & \text{otherwise}\end{cases}. \tag{5.4}$$

Using overbars to denote cross-sectional averages for fixed current and birth dates, we obtain the following expression for the difference in the consumption of two successive cohorts when $t-s$ is even:

$$\begin{aligned}\overline{\log c_{t,s}}-\overline{\log c_{t,s-1}} &= -\frac{1}{\gamma}\log\beta+\overline{\log c_{s,s}}+\frac{1}{\gamma}\log\xi_s-\overline{\log c_{s-1,s-1}}-\frac{1}{\gamma}\log\xi_{s-1}\\ &\quad -\frac{1-q}{\gamma}\left[\log\frac{\xi_t}{\xi_{t-1}}-\log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right].\end{aligned}$$

When $t-s$ is odd, we obtain

$$\begin{aligned}\overline{\log c_{t,s}} - \overline{\log c_{t,s-1}} &= -\frac{1}{\gamma}\log\beta + \overline{\log c_{s,s}} + \frac{1}{\gamma}\log\xi_s - \overline{\log c_{s-1,s-1}} - \frac{1}{\gamma}\log\xi_{s-1} \\ &\quad + \frac{(1-q)}{\gamma}\left[\log\left(\frac{\xi_t}{\xi_{t-1}}\right) - \log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right].\end{aligned}$$

OLS estimates of first differences in cohorts are formed by averaging $\overline{\log c_{t,s}} - \overline{\log c_{t,s-1}}$ across $t = 1...T$.[6] Letting T_1 be the number of even observation and T_2 the number of odd observations, we obtain

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left(\overline{\log c_{t,s}} - \overline{\log c_{t,s-1}}\right) \\ &= -\frac{1}{\gamma}\log\beta + \overline{\log c_{s,s}} + \frac{1}{\gamma}\log\xi_s - \overline{\log c_{s-1,s-1}} - \frac{1}{\gamma}\log\xi_{s-1} \\ &\quad -\frac{(1-q)}{\gamma}\left[\frac{1}{T_1}\sum_{t:\frac{t}{2}\in\{1,2...\}}\log\left(\frac{\xi_t}{\xi_{t-1}}\right) - \frac{1}{T_1}\sum_{t:\frac{t}{2}\in\{1,2...\}}\log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right] \\ &\quad +\frac{(1-q)}{\gamma}\left[\frac{1}{T_2}\sum_{t:\frac{t}{2}\notin\{1,2...\}}\log\left(\frac{\xi_t}{\xi_{t-1}}\right) - \frac{1}{T_2}\sum_{t:\frac{t}{2}\notin\{1,2...\}}\log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)\right].\end{aligned} \tag{5.5}$$

As T becomes large, each of the the last two lines approaches zero. Indeed, a law of large numbers implies that $\frac{1}{T_1}\sum_{t:\frac{t}{2}\in\{1,2...\}}\log\left(\frac{\xi_t}{\xi_{t-1}}\right)$ converges in probability to $E\log\left(\frac{\xi_t}{\xi_{t-1}}\right)$. Similarly, a law of large numbers implies that $\frac{1}{T_1}\sum_{t:\frac{t}{2}\in\{1,2...\}}\log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right) = E\log E_{t-1}\left(\frac{\xi_t}{\xi_{t-1}}\right)$. The same obtains for the case where the summation runs over $t : \frac{t}{2} \notin \{0,1,2...\}$. Therefore the terms inside square brackets become asymptotically equal, so that the last two lines on the right-hand side of (5.5) disappear.

Interestingly, the terms on the first line of equation (5.5) are the same, whether there is inattention or not. Hence, all the conclusions that we derive in the paper about the permanent components of cohort effects and their relationships to the stochastic discount factor remain intact.

[6] As we explain in the text such cohort differences are exactly identified when we only include time and cohort effects along with a constant in the regression. When we additionally include age dummies, then the differences in cohort effects are identified up to a constant. Since our results depend on the variances and covariances of differences in cohort effects, the fact that we cannot identify the constant is inconsequential.

We conclude with a clarifying remark. The above derivations are not meant to convey that inattention will not impact the stochastic discount factor. In general it will. Our claim is that the permanent components of cohort effects will still reveal $\log(\xi_t) + \gamma \log(C_t)$, which in the model is equal to the displacement factor. Hence, the conclusion of the paper that changes in the permanent component of cohort effects combined with aggregate consumption growth provide an accurate description of the stochastic discount factor remains valid in the presence of inattention.

6 Declining Prices of Intermediate Goods

Our model makes three assumptions: a) the cost of producing an intermediate good is constant, i.e., it takes one unit of labor to produce a single good, and b) each blueprint is produced by a firm that holds monopoly rights to its production forever. Coupled with the simple Dixit-Stiglitz aggregator we use, these assumptions lead to the usual "constant mark-up" rule:

$$p_{j,t} = \frac{w_t}{\alpha}, \tag{6.1}$$

which is common to many models of monopolistic competition. An implication of equation (6.1) is that the relative prices of any intermediate goods j and j' remain the same no matter how much time has elapsed since their introduction. This implies that in our model displacement of incumbents' profits operates exclusively through reductions in the quantities they produce rather than the prices they charge. Even though this distinction between price and quantity effects on incumbent profits is inconsequential for the asset pricing conclusions of our paper, we note that in reality it is not exclusively the quantities, but also the prices of existing goods that decline in response to innovation.[7]

There are at least two ways to extend the model so as to account for declining prices of intermediate goods. We could either assume declining costs of production (say, due to

[7]See, e.g., Gort and Klepper (1982), Klepper and Graddy (1990).

learning by doing) or — more realistically — allow for the entry of "immitators", who produce perfect (or at least close) substitutes to existing intermediate goods.

The first modification (declining costs of production) is fairly straightforward to incorporate into the baseline model. Specifically, letting $\varphi \in (0, 1)$ and assuming that it takes $\left(\frac{j}{A_t}\right)^{\varphi}$ units of labor to produce intermediate good j, equation (6.1) becomes

$$p_{j,t} = \left(\frac{j}{A_t}\right)^{\varphi} \frac{w_t}{\alpha}.$$

With this simple modification, the relative price of a fixed intermediate good $p_{j,t}$ compared to the "high tech" intermediate good $p_{A_t,t}$ is declining over time, since $\frac{p_{j,t}}{p_{A_t,t}} = \left(\frac{j}{A_t}\right)^{\varphi}$. (The same conclusion holds if we compute the relative price of a fixed intermediate good j with respect to the price index of all intermediate goods). Furthermore, as long as φ is small enough, all the conclusions of our model with respect to displacement etc. continue to hold.

An arguably more realistic alternative to obtain declining intermediate prices is to assume entry of firms that produce perfect substitutes to intermediate good j. Such an extension can be easily accommodated within our framework. To illustrate in the simplest possible way how to extend the model in such a direction, we assume that, in addition to "innovators" who introduce new blueprints, there exist also "imitators" who introduce perfect substitutes to existing blueprints. For illustration, suppose that such imitators introduce a new perfect substitute to intermediate good j with probability p each period. Then the final-goods firm production function becomes

$$Y_t = Z_t \left(L_t^F\right)^{1-\alpha} \int_0^{A_t} \omega_{j,t} \left(\sum_{n_j = 1..N_j} x_{j,t}^{n_j} \right)^{\alpha} dj,$$

where $x_{j,t}^{n_j}$ is the amount of intermediate goods purchased by producer n_j in the market for blueprint j, and $N_j - 1$ is the total number of imitators that have entered the market since its inception by the initial innovator. Assuming that increased competition reduces markups the prevailing price in market j will gradually decline from $\frac{w_t}{\alpha}$ to w_t as more and more imitators enter. Moreover, the younger the market, the closer the price will be to $\frac{w_t}{\alpha}$. By contrast, older markets will have experienced more entry of imitators, and hence the price will be closer to the marginal cost of production w_t.

One can extend this idea further by introducing endogenous entry into incipient industries, refinements in the production of existing goods, exit, etc.. Such a model would address the evolution of market structure within existing industries, while displacement risk — i.e., the arrival of new industries — would present a systematic adverse demand shock across all existing industries.

7 Implications for Optimal Risk Sharing between Agents

In this section we investigate in greater depth how portfolio choices of agents with heterogeneous financial-to-total-wealth ratios lead to optimal risk sharing. Specifically, we compute the equilibrium exposures of the financial and human capital components of aggregate wealth to the fundamental shocks in the model. Using these exposures, we then infer the heterogeneous portfolio holdings (value or growth tilt) of agents with different human-to-total capital ratios.

To start, we note that a worker's financial wealth can be determined at any point in time from the intertemporal budget constraint. To economize on notation, we start by letting

$$\begin{aligned} B_1 &\equiv E_t \sum_{t'=t}^{\infty} \left(\frac{\xi_{t'}}{\xi_t}\right) (1-\lambda)^{(t'-t)} \left(\frac{c_{t',s}}{c_{t,s}}\right), \\ B_2 &\equiv E_t \sum_{t'=t}^{\infty} \left(\frac{\xi_{t'}}{\xi_t}\right) \left[(1-\lambda)(1+\delta)\right]^{(t'-t)} \left(\frac{q_{t',s} w_{t'}}{q_{t,s} w_t}\right). \end{aligned}$$

We note that B_1 and B_2 are constants in our model. Using the definitions of B_1 and B_2, the intertemporal budget constraint of a worker at time t can be expressed as

$$B_1 c_{t,s} = W_{t,s} + B_2 y_{t,s}, \tag{7.1}$$

where $y_{t,s} \equiv \overline{h} w_t q_{t,s} (1+\delta)^{t-s}$ is the agent's labor income at time t. Equation (7.1) states that a worker's net present value of consumption ($B_1 c_{t,s}$) is equal to her *total* wealth, i.e. the sum of her financial wealth ($W_{t,s}$) and the net present value of her earnings $B_2 y_{t.}$

An implication of (7.1) is that

$$\Delta \log (W_{t+1,s} + A_2 y_{t+1,s}) = \Delta \log c_{t+1,s}, \tag{7.2}$$

where $\Delta x_{t+1} = x_{t+1} - x_t$. Importantly, the fact that agents of different cohorts share risk frictionlessly implies that both the left- and the right-hand sides of (7.2) are independent of s. (This is an implication of equation 21 in the paper). Focusing on agents with $W_{t,s} > 0$, we next approximate

$$\Delta \log \left(W_{t+1,s} + A_2 y_{t+1,s}\right) \approx (1 - \vartheta_t) \frac{\Delta W_{t+1,s}}{W_{t,s}} + \vartheta_t \Delta \log \left(y_{t+1,s}\right), \tag{7.3}$$

where $\vartheta_t \equiv \left(\frac{A_2 y_{t+1,s}}{W_{t+1,s}+A_2 y_{t+1,s}}\right)$ is the fraction of total wealth due to the net present value of labor income. Using the definition of y_t inside (7.3) gives

$$\Delta \log \left(y_{t+1,s}\right) = \left[(1+\delta) - (\rho + \alpha - 1)\, u_{t+1} + \varepsilon_{t+1}\right]. \tag{7.4}$$

Furthermore, the model implies that

$$\Delta \log c_{t+1,s} \approx \varepsilon_{t+1} + \left[(1-\alpha) + \frac{\upsilon'(0)}{\upsilon(0)}\right] u_{t+1}, \tag{7.5}$$

where the function $\upsilon(\cdot)$ is given in equation (25) of the paper. Combining (7.2), (7.3), (7.4) and (7.5) gives

$$\frac{\Delta W_{t+1,s}}{W_{t,s}} \approx \left[(1-\alpha) + \frac{1}{1-\vartheta_t}\left(\frac{\upsilon'(0)}{\upsilon(0)} + \vartheta_t \rho\right)\right] u_{t+1} + \varepsilon_{t+1} - \frac{\vartheta_t}{1-\vartheta_t}(1-\delta). \tag{7.6}$$

The term inside square brackets in equation (7.6) is of particular interest for our purposes. It captures the proportional change in an agent's financial wealth in response to a shock to u_{t+1}. More importantly, this term allows us to infer whether the investor has a "growth" tilt or a "value" tilt in her portfolio. For instance, if the term inside square brackets is small in absolute value, this means that the investor has bought enough growth stocks (and/or shorted value stocks) so as to approximately immunize the exposure of her financial wealth to displacement risk. By contrast, for investors who choose a high exposure to value stocks, the term inside square brackets is negative and large in absolute value.

To determine whether agents who have a high human-to-total wealth ratio (ϑ_t) choose more or less exposure to displacement risk, we first note that the term inside square brackets in equation (7.6) is negative, since in our calibration $(1-\alpha) + \frac{\upsilon'(0)}{\upsilon(0)} + \rho < 0$. Furthermore, a direct differentiation with respect to ϑ_t reveals that the term is declining in ϑ_t.[8] Accordingly, agents with a high human-to-total wealth (ϑ_t) ratio choose portfolios that make their financial wealth more negatively exposed to displacement risk (value tilt). In contrast, agents with a low ϑ_t have a less negative exposure to displacement risk (u_{t+1}), i.e., they have a tilt towards growth stocks in their portfolio.

8 Education Choice, the Skill Premium, and Cross-Sectional Wage Dispersion

Since the model's focus is on asset pricing, in its baseline version we abstract from issues related to education choice. As a result, the baseline model is silent about the skill premium and cross-sectional wage dispersion. Here we enrich the model to allow for education choice in a stylized way. We show that allowing for such choice leads to cross-sectional dispersion in the wages within a given cohort. The cross-sectional dispersion is increasing in the size of the technological shock u_s at the time when the cohort enters the workforce.

For simplicity, we ignore external habit formation ($\psi = 1$) and make such functional-form choices that the model's implications for aggregate prices do not change. Specifically, suppose that at the time of her "birth" s, a worker i chooses (once and for all) her level of educational attainment e_s^i. The benefit of education is that it raises a worker's productivity. Specifically, a worker's total supply of efficiency units at time t is given by $e_s^i q_{t,s}$, where $q_{t,s}$ captures the efficiency units of labor, as specified in the paper.

Assume that education involves only non-pecuniary costs that are lower for agents with

[8]Direct differentiation of the term with respect to ϑ_t, along with the fact that $\frac{\upsilon'(0)}{\upsilon(0)} + \rho < -(1-\alpha) < 0$, implies that

$$\left[(1-\alpha) + \frac{1}{1-\vartheta_t}\left(\frac{\upsilon'(0)}{\upsilon(0)} + \vartheta_t \rho\right)\right]' = \frac{1}{(1-\vartheta_t)^2}\left[\frac{\upsilon'(0)}{\upsilon(0)} + \rho\right] < 0.$$

stronger scholastic skills. Specifically, the agent maximizes

$$E_s \sum_{s=t}^{\infty} [\beta (1-\lambda)]^{t-s} \frac{\left[\left(1-\frac{e_s^i}{2z_s^i}\right) c_{t,s}^i\right]^{1-\gamma}}{1-\gamma}, \tag{8.1}$$

where z_s^i is an idiosyncratic, time-invariant, person-specific shock, meant to capture scholastic aptitudes. In the specification (8.1) we follow the lead of the seminal RBC paper by King et al. (1988) and assume that the non-pecuniary costs of education are multiplicatively separable in a term involving educational attainment e_s^i and the agent's consumption. As King, Plosser and Rebello show in the closely related context of the choice between leisure and work, the multiplicative separability of the utility specification is necessary in order to ensure that e_s^i remains stationary as the economy grows — "balanced growth".

We next show that as long as we normalize the distribution of z_s^i judiciously, all equilibrium prices remain the same as in the original version of the model. The choice of educational attainment e_s^i scales agent i's endowment in each state and at all times by a factor of e_s^i. Due to the iso-elastic form of the utility function, this in turn scales the agent's expected life-time utility of consumption by a factor of $(e_s^i)^{1-\gamma}$. According to (8.1), the non-pecuniary cost of education further scales the agent's life-time utility by a factor of $(1-e_s^i/(2z_s^i))^{1-\gamma}$. Thus, the net effect of choosing the level of education e_s^i is to scale the life-time utility of the agent by a factor of

$$\left[\left(1-\frac{e_s^i}{2z_s^i}\right) e_s^i\right]^{1-\gamma}, \tag{8.2}$$

and hence the optimal choice of educational attainment is

$$e_s^i = z_s^i. \tag{8.3}$$

Equation (8.3) states that workers with higher scholastic aptitude z_s^i choose more education. Furthermore, if we normalize $E\left(z_s^i\right) = 1$, the cross-sectional mean $E\left(e_s^i\right)$ is unity at all points in time. As a result, the aggregate value of labor income of the cohort born at time s is identical to the expression given in the baseline model. Since all prices (aggregate wage, stochastic discount factor, etc.) are only affected by the aggregate value of labor income accruing to each cohort, all equilibrium prices are unaffected by the introduction of education choice.

Even though equilibrium prices are unchanged relative to the model without education choice, the allocation of consumption is. Now the model features cross-sectional dispersion in labor income, simply because agents with higher education earn more labor income. Interestingly, taking two cohorts s and s', and computing the cross-sectional variance of labor income we find that

$$\frac{Var\left(e_s^i q_{t,s} w_t\right)}{Var\left(e_{s-1}^i q_{t,s-1} w_t\right)} = \frac{q_{s,s}^2 e^{2\rho u_s}}{q_{s-1,s-1}^2}. \tag{8.4}$$

Since $q_{s,s} > 0$ is an increasing function of u_s, equation (8.4) implies that $u_s > u_{s-1}$ leads to $Var\left(e_s^i q_{t,s} w_t\right) > Var\left(e_{s-1}^i q_{t,s-1} w_t\right)$. Alternatively phrased, cohort s, having experienced higher technological advancement compared to its preceding cohort, also exhibits a higher cross-sectional dispersion of wages.

The above calculations show that the model is consistent with the empirical observations in Attanasio and Davis (1996) on cohort-specific skill-premia, even if educational choice is purely meritocratic, i.e., exclusively determined by agents' innate scholastic aptitudes $\left(e^i\right)$. The above qualitative conclusions would not change if we introduced meritocratic educational choice by allowing education to have monetary costs with workers who can borrow against the increase of their labor income due to education. (In such a model, an agents' labor income would have to be modelled as $z^i e_s^i q_{t,s}$). However, such a model would require specifying an educational sector that uses resources, and the model would become more involved.

9 The Value Premium and the IT Revolution

According to the model, the realized return differential between value and growth stocks should decline during periods of increased displacement activity. This implication of the model appears to be broadly consistent with the behavior of the value premium during the late eighties and nineties, two periods associated with the adoption of IT technology.

Figure 1 plots the 10-year moving average of the annual value spread for the US, so as to isolate low frequency movements of the value spread. The figure shows a significant decline in the value spread in the eighties and particularly in the nineties, when the 10-year moving average actually turns negative. Accepting the popular view that the eighties and especially the nineties were periods of accelerated technological growth, this evidence is in line with the theory put forth in the paper.

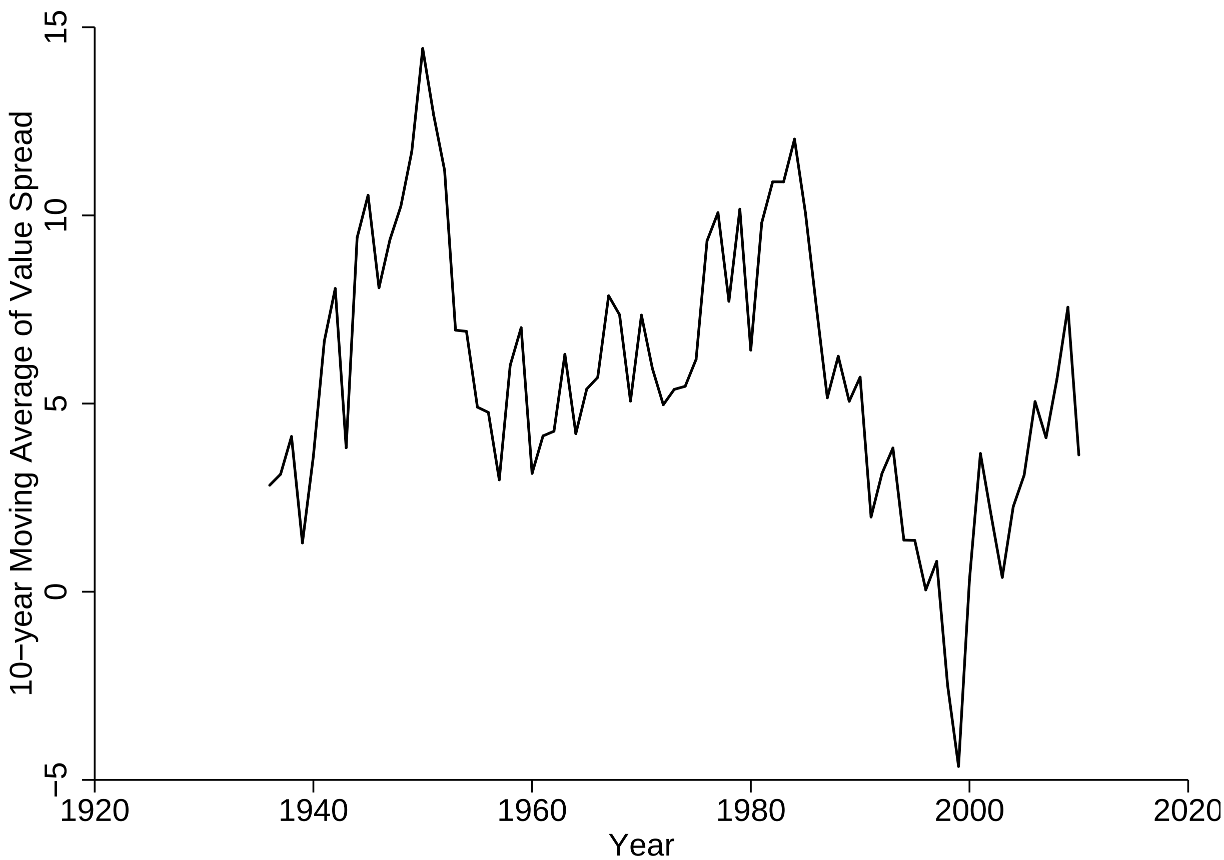

Figure 1: 10-year moving average of the value spread. To enable comparisons with the international value and growth portfolios of Fama and French, we define the value spread as the difference in returns between stocks in the top and bottom 30 percent when sorted by book-to-market ratio.

The international data provides further evidence in this direction. If the IT revolution was a global phenomenon, then we should see declines in the value premium across the world.

The advantage of international data is that one can exploit the panel dimension to test this hypothesis. However, a limitation of international data is that the only readily available data on the value spread start in 1975 (or later) for most countries. Because of this limitation we cannot produce an analog of figure 1 for each country and "eyeball" the onset of the decline of the value premium. Therefore, we adopt a regression approach. We choose a period during which the IT revolution is most likely to be spreading across all countries of the world. Both because the US is likely to have been one of the early adopters, and because the time-series dimension of our sample is rather short, we choose the years 1994-1999 for our baseline estimation. The first row of table 1 reports results of regressing the annual value spread on country fixed effects and a dummy variable taking the value one during 1994-1999. We find that the value spread is on average -8.7 percent lower across the world during this period. To put this in perspective, the average value spread across all year-country observations in our sample is 4.67 percent. As a robustness check, the second row of the table reports results when the indicator variable takes the value one in the longer period 1985-1999. The sign remains the same, but the magnitude and the statistical significance of the results becomes smaller. (In particular, when standard errors are clustered by year, then the one-sided test that the coefficient is non-positive is marginally significant at the 5 percent level, but the two-sided test is not.) In non-reported results we experimented with more starting dates between 1985 and 1994. The coefficients were negative, and the results became stronger as the starting date approached 1994. Taken together the international evidence supports the notion that the value premium declined during the IT revolution, and especially during the latter half of the nineties.

10 Robustness Checks with respect to the Parameters in the Baseline Model

Table 2 reports results from some simple robustness exercises. The column titled $\psi = 1$ helps isolate the effect of habit formation. This column shows how results change in the case where

	Indicator (1994-1999)	Indicator (1985-1999)
Coefficient	-8.70	-6.40
t-stat	(-3.32)	(-3.15)
t-stat (clustered)	(-2.45)	(-1.67)
Country Dummies	Yes	Yes
Observations	734	734

Table 1: The value spread during periods of technological expansion: International evidence. We use an international panel of annual return differentials between value portfolios (top-30th percentile of the book to market distribution) and growth portfolios (bottom-30th percentile of the book to market distribution) provided on the website of Kenneth French. We used all countries (Austria, Australia, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zeland, Norway, Singapore, Spain, Sweden, Switzerland, UK, and US) except for Malaysia, because the available sample for that country was too short for our purposes (1994-2001). The rest of the panel is unbalanced, with most countries having sample periods starting in 1975. The first line reports the coefficients on a dummy variable taking the value one in the period 1994-1999 (first row). The second row reports the results for a dummy taking the value one in the years 1985-1999. Below each coefficient, we report conventional t-stats and (robust) t-stats clustered by year. Country fixed-effects are included in both regressions, but not reported.

agents have standard constant relative risk aversion preferences ($\gamma = 10$). Comparing this table to Table 5 in the paper, it is apparent that the absence of habit formation increases slightly both the equity and the value premium. However, this comes at the cost of also increasing the riskless rate[9] and as a result all the earnings-to-price ratios. The next column ($\kappa = 0.7$) reduces κ to 0.7 while keeping the rest of the parameters unchanged. Recall that κ reflects the fraction of new blueprints accruing to new firms owned by arriving agents,

[9]Inspection of equation (24) helps with the intuition behind this result: As ψ increases from 0 to 1, the exponent of $\frac{Y_{t+1}}{Y_t}$ decreases from -1 to $-\gamma$, resulting in a higher volatility of the pricing kernel, but also a stronger negative drift of the (log) stochastic discount factor.

	Data	$\psi = 1$	$\kappa = 0.7$ baseline	$\kappa = 0.7$ modified
Aggregate (log) consumption growth rate	0.017	0.017	0.017	0.017
Aggregate (log) consumption volatility	0.033	0.032	0.032	0.033
Riskless rate	0.010	0.055	0.024	0.017
Equity premium	0.061	0.046	0.027	0.032
Aggregate earnings / price	0.075	0.144	0.091	0.106
Dividend volatility	0.112	0.108	0.081	0.107
Correl. (divid. growth, cons.growth)	0.2	0.14	0.25	0.12
Std $(\Delta\alpha_s^{\text{perm}})$	0.023	0.029	0.026	0.027
$\frac{cov\left(\Delta\alpha_s^{\text{perm}},\text{ growth-value }\right)}{var\left(\Delta\alpha_s^{\text{perm}}\right)}$	3.92	3.865	5.257	6.750
Std $(\Delta w_s^{\text{perm}})$	0.022	0.021	0.021	0.023
Earnings / price 90th Perc.	0.120	0.153	0.110	0.140
Earnings / price 10th Perc.	0.04	0.066	0.036	0.042
Average value premium	0.065	0.067	0.065	0.074
Std (growth-value)	0.127	0.116	0.137	0.182
$\mathrm{E}\left(R^a - R^o\right)$		0.118	0.096	0.106

Table 2: Robustness Checks. The columns titled $\psi = 1$ and $\kappa = 0.7$ display results when the parameters ψ and κ are set equal to 1 and 0.7 respectively, while the rest of the parameters are kept at their baseline values ($\gamma = 10$). The last column displays results assuming that $\kappa = 0.7$, $\chi = 5$, $\nu = 0.06$, $\rho = 0.8$, and the rest of the parameters are kept at their baseline values.

while $(1-\kappa)$ accrues to existing agents. In the baseline case scenario we choose $\kappa = 0.9$. We consider this a plausible value for the following reason: In a fully specified endogenous-innovation model with capital, where factors of production in the innovation sector are compensated for their marginal product, κ would capture the share of human capital, labor and skill in the innovation process (as opposed to the share of capital operated by pre-existing firms). Assuming that education, entrepreneurial skill, and human effort are the most important scarce factors in the innovation process, one would expect κ to be close to 1. (For instance, in the seminal Romer model, labor is the exclusive factor of production in the development of new ideas). However, to examine the robustness of the results to this assumption, we also examine what happens when we choose these income shares to be similar to aggregate income shares in NIPA data. To that end we choose $\kappa = 0.7$. The next to last column reports results when all other parameters are kept at their base values, while the last column reports what happens when the rest of the parameters are also modified in order to match the volatility of dividends. As can be seen, even though the results are slightly weaker when $\kappa = 0.7$, the model retains its power to explain a large fraction of the observed moments in the data.

References

Abel, A. B., J. C. Eberly, and S. Panageas (2010). Optimal inattention to the stock market with information costs and transactions costs. mimeo. Wharton, Kellogg, and Chicago Booth.

Attanasio, O. and S. J. Davis (1996). Relative wage movements and the distribution of consumption. *Journal of Political Economy 104*(6), 1227 – 1262.

Duffie, D. and T.-s. Sun (1990). Transactions costs and portfolio choice in a discrete-continuous-time setting. *Journal of Economic Dynamics and Control 14*(1), 35–51.

Gabaix, X. and D. Laibson (2002). *The 6D Bias and the Equity-Premium Puzzle*, pp. 257–312. NBER macroeconomics annual 2001, Bernanke, Ben S. and Rogoff, Kenneth, eds. Volume 16. Cambridge and London: MIT Press.

Gort, M. and S. Klepper (1982). Time paths in the diffusion of product innovations. *Economic Journal 92*(367), 630 – 653.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988). Production, growth and business cycles: I. the basic neoclassical model. *Journal of Monetary Economics 21*(2/3), 195–232.

Klepper, S. and E. Graddy (1990). The evolution of new industries and the determinants of market structure. *RAND Journal of Economics 21*(1), 27 – 44.