

Online Appendix

for “Production Networks and Stock Returns: The Role of Vertical Creative Destruction”

Michael Gofman^{*}

Gill Segal^{*}

Youchang Wu[†]

January 19, 2020

Abstract

This document supplements the paper “Production Production Networks and Stock Returns: The Role of Vertical Creative Destruction”. In this appendix we: (1) Provide sensitivity analysis to the model’s parameters; (2) Describe the augmented model with monopolistic competition and the construction of the supply-chain competition measure; (3) Show higher input cyclicalities for upper layers using BLS data; (4) Discuss several model assumptions and extensions; (5) Describe the details regarding the vertical position computation using BEA I-O tables and using Compustat Segment database; (6) Discuss whether the TMB spread is a priced factor; (7) Rule out mechanical results and a few other alternative explanations for the TMB spread. We also present a collection of supplemental tables.

Contents

OA.1	Sensitivity Analysis.....	OA - p.2
OA.2	An Augmented Model with Monopolistic Competition	OA - p.5
OA.3	A Measure of Supply Chain Competition	OA - p.7
OA.4	Cyclicalities of Input Prices Using BLS FD-ID Database.....	OA - p.8
OA.5	Model Assumptions and Extensions.....	OA - p.10
OA.5.1	Goods transmitted in supply-chains	OA - p.10
OA.5.2	Ex-ante Layer Heterogeneity	OA - p.11
OA.5.3	Alternative Network Specifications	OA - p.11
OA.5.4	Transition between Layers: A Quantitative Assessment	OA - p.14
OA.6	Vertical Position Computation Using Input-Output Tables.....	OA - p.16

^{*}University of Rochester, E-Mail: Michael.Gofman@simon.rochester.edu.

^{*}Corresponding Author, University of North Carolina at Chapel Hill, E-Mail: Gill_Segal@kenan-flagler.unc.edu, Address: McColl 4113, CB 3490 Chapel Hill, NC 27599-3490, Phone: +1 (919) 962-3854.

[†]University of Oregon, E-Mail: ywu2@uoregon.edu.

OA.7	Vertical Position Computation Compustat Segment Data	OA - p.19
OA.8	TMB: Is it a Priced Risk Factor?	OA - p.21
OA.9	Ruling Out Mechanical Results	OA - p.25
OA.10	Two Additional Alternative Explanations	OA - p.28
OA.11	Supplemental Tables	OA - p.30

OA.1 Sensitivity Analysis

In this subsection we demonstrate that the TMB spread is *qualitatively* robust to most calibration parameters. We point out the quantitative importance of the parameter choices.

Table OA.1: **Sensitivity Analysis to Model Parameters**

	(1) Benchmark Calibration	(2) IES = 1.1	(3) IES = 0.9	(4) LR-Vol=0; SR-Vol=1.8%	(5) $\delta = 0.05$	(6) No Adj. Costs
Mean Excess Returns (%)						
Layer 5	16.07	3.84	1.20	0.87	17.27	19.12
Layer 4	12.01	3.43	1.34	0.86	14.52	5.71
Layer 3	9.42	3.17	1.40	0.84	10.39	4.23
Layer 2	7.15	2.91	1.50	0.82	7.32	3.52
Layer 1	5.23	2.64	1.64	0.78	5.02	2.61
Layer 0	3.59	2.36	1.86	0.72	3.29	1.43
Spread (5-0)	12.49	1.49	-0.66	0.15	13.98	17.69
Mean Returns (%)						
Market Portfolio	4.13	2.40	1.82	0.75	3.85	1.94
Risk-free rate	1.02	3.57	5.12	4.40	1.13	0.67
Consumption Growth						
Mean (%)	1.94	1.79	1.98	1.94	2.01	1.99
Standard deviation (%)	1.74	1.94	1.98	1.74	1.73	1.80

The table presents summary model results using different calibrations. The left most column shows the variable of interest. Column (1) presents results from the model under the benchmark calibration. Columns (2) - (6) present results from the model calibrated using the same parameters as in the benchmark case, other than the parameter(s) specified right below the column number. In Column (2) the IES parameter ψ is set to 1.1. In Column (3) the IES parameter is set to 0.9. In Column (4) the long-run volatility is set to zero, and the short-run volatility is raised to 1.8%. In Column (5) the capital depreciation rate δ is reduced to 5%. In Column (6) the capital adjustment cost parameter ϕ is set to 1 (no adjustment costs). All model implied moments are based on a simulated population path of length 100,000 periods.

The Role of IES. Column (1) of Table OA.1 shows the summary model statistics for a model with a similar calibration to the benchmark case, except that the IES parameter, ψ , is reduced to 1.1 (still above unity). The drop in the elasticity of intertemporal substitution parameter raises the level of the risk-free rate and drops the level of the equity premium compared to the benchmark

case. Importantly, the spread in the returns between layers 5 and 0 is still positive, but diminished in magnitude to 1.49%. Intuitively, a lower IES implies that firms invest less in response to long-run productivity shocks, which attenuates their productivity betas. Column (2) of Table OA.1 shows the results when the IES parameter is dropped below the unity threshold to 0.9. In that case, the spread between the layer 5 and layer 0 excess returns turns negative. When the IES is less than unity, the income effect dominates the substitution effect. In response to a productivity news shock, the household desires to increase consumption strongly. The productivity news shock acts as a positive demand shock for the final goods, but less so for intermediary goods (that is, saving or capital goods). This positive demand shock diverts resources to firms at bottom layers. The capital installation costs become larger for these bottom layer firms, increasing their Tobin's Q and making them more sensitive to productivity shocks.

The Role of Long-Run Productivity. In Column (3) of Table OA.1 we report the results when we shut down the long-run productivity news shocks ($\phi_x = 0$). To keep consumption growth volatility constant at the benchmark case level, we simultaneously raise the short-run production volatility σ_z to 1.8%. In this case, consumption growth volatility remains 1.74% per annum. However, the equity premium is diminished to only 0.75% per annum. Similarly, the spread between layer 5 and layer 0 remains positive, but it is only 0.15%. As shown in Croce (2014), short-run productivity shocks do not induce firms to invest largely enough, which implies positive yet low productivity betas. Introducing a persistent component to the growth of aggregate productivity, as in the benchmark case, causes firms to react more strongly to technology news (so long as the substitution effect dominates), and amplified betas.

The Role of Depreciation Rate. Creative destruction in the model puts a downward pressure on the valuation of installed capital stock of customer firms. Thus, the creative destruction channel has a stronger quantitative bite when capital depreciates more slowly, and the stock of existing capital has a persistent impact on future periods. Put differently, with a lower depreciation rate the relative importance of assets-in-place out of firm value rises, which is the component that is subject to the creative destruction mechanism. In line with this intuition, column (3) of Table OA.1 shows that the spread is amplified when the depreciation rate is reduced from 10% in the benchmark case to 5%. The reverse is also true: with higher depreciation rates, the importance of the capital stock is diminished, and the spread gets smaller.

In fact, we can show that in the extreme case in which depreciation is full and adjustment cost is zero, there is no creative destruction on pre-installed capital, and no dispersion in firms' assets-in-place riskiness. To see this, note that by constant returns to scale, one can write the value of a firm in layer $j \in \{0..N-1\}$ as: $V_{j,t} = d_{j,t} + Q_{j,t}k_{j,t+1}$. Plugging the expression for the dividend and capital's law of motion yields:

$$\begin{aligned} V_{j,t} &= Y_{j,t} - W_t n_{j,t} - i_{j,t} k_{j,t} P_{j+1,t} + Q_t((1-\delta)k_{j,t} + i_{j,t} k_{j,t}) \\ &= d_{j,t}^* + Q_{j,t} k_{j,t} (1-\delta), \end{aligned}$$

where the second equality follows from the fact that $Q_{j,t} = P_{j+1,t}$ in the absence of adjustment costs, and $d_{j,t}^* \equiv \max_{n_{j,t}} Y_{j,t}(n_{j,t}) - W_t n_{j,t}$. We define the dividend $d_{j,t}^*$ as the operating profit of layer j at time t (that is, output net of wage payments). Let $V_{j,t}^{ex} = V_{j,t} - d_{j,t}^*$ be the ex-dividend valuation of layer j . Clearly, $V_{j,t}^{ex} = Q_{j,t} k_{j,t} (1-\delta)$. When $\delta = 1$, the ex-dividend firm value is zero, suggesting that $\beta_j^{AP} = \partial V_{j,t}^{ex} / \partial \varepsilon_{z,t} = 0$ for all j .

As long as the depreciation is not full (i.e., $\delta = 1$), creative destruction is at work (as Q multiplies a positive stock), and the TMB spread is positive. Understanding how depreciation affects vertical creative destruction is important given the fact that in principle, some suppliers sell intermediate goods, which can be viewed as a form of capital that depreciates faster. To see how the results may change if we interpret our capital good as an inventory, we raise the depreciation rate to 20%. This value is consistent with the findings of Richardson (1995) and Jones and Tuzel (2013) that noninterest inventory carrying costs, which are equivalent to depreciation, vary from 19% to 43% per annum. We find that the (local) sensitivity of the spread to the δ parameter is relatively small. Under a depreciation rate of 20% the spread drops by a small amount to 12.00% per annum.

The Role of Capital Adjustment Costs. Column (6) of Table OA.1 presents the results when there are no capital adjustment costs ($\phi = 1$). Without adjustment costs, the spread between layers 5 and 0 becomes even larger. The creative destruction channel, which is the primary force behind our benchmark result, is amplified in the *absence* of adjustment costs. Notice that without adjustment costs, the TMB spread is higher than the benchmark calibration (17.69% versus 12.49%). The spread is now too large compared to the data. In addition, the average market excess return drops to only 1.94%, and the excess return for the firm in the bottom layers is counterfactually low (about 1.4%).

The intuition behind this result is that in the absence of adjustment costs, a larger fraction

of firms' value comes from assets-in-place. Since assets-in-place is the component that is subject to the creative destruction force, the spread is amplified. To see that, note that in a perfect competition model, firms earn rents from positive productivity shocks (that is, have higher growth option component) due to the existence of capital adjustment costs. The adjustment costs are in essence an installation costs that create a wedge between the price of new uninstalled capital goods and the marginal value of installed capital (Tobin's Q). This wedge boosts Tobin's Q and generates rents. To formalize this idea, we define the value of a firm's assets-in-place as the value of all future dividends resulting from the existing capital stock, which only depreciates over time:

$$V_{j,AIP,t} = \max_{n_{j,t}} P_{j,t} Z_t k_{j,t}^\alpha n_{j,t}^{1-\alpha} - w_t n_{j,t} + (1 - \delta) E[M_{t,t+1} V_{j,AIP,t+1}] \quad \forall j \in \{0..N\}.$$

We then define the weight of assets-in-place for a firm j as the ratio between $V_{j,AIP,t}$ and the total firm value $V_{j,t}$. In the benchmark calibration the average weight of assets-in-place is 52% across the layers. As the capital adjustment cost parameter ϕ drops to 1 in Column (6), the ratio becomes 100% in the limit. In other words, in column (6) firms look more like "value" firms compared to the benchmark case, which enhances the spread.¹

OA.2 An Augmented Model with Monopolistic Competition

In this section, we present an augmented model that accommodates monopolistic competition.

Aggregate productivity has the same dynamics as those described in Section 5.1. The household side of the economy is identical to that described in Section 3.2. Unlike the perfect competition model, we now assume that each production layer $j \in \{0..N\}$ is populated by a mass of differentiated intermediate good producers, indexed by $m \in [0, 1]$. The output of an intermediate good producer in layer j at time t of variety m is denoted by $y_{j,t}(m)$. Its output price is $p_{j,t}(m)$.

Aggregators. In each layer j , an aggregator converts the layer's intermediate goods into a final composite layer good, $Y_{j,t}$, using a CES production function:

$$Y_{j,t} = \left[\int_0^1 y_{j,t}(m)^{\frac{\mu_j - 1}{\mu_j}} dm \right]^{\frac{\mu_j}{\mu_j - 1}}, \quad (\text{OA.1})$$

when $\mu_j \rightarrow \infty$, the intermediate good producers of the j -th layer face perfect competition. For any finite μ_j , the intermediate good producers are not perfect substitutes, and they possess some amount of monopolistic power.

¹The intuition for the decline in the equity premium in column (6) is straightforward. By excluding adjustment costs, productivity shocks are absorbed in how much the firms invest (that is, in quantities), as opposed to being absorbed in prices. This allows the firms to smooth their dividends more easily, making their valuations less volatile. As a result, the equity risk premium drops.

The j^{th} layer aggregator faces perfect competition in the product market. It solves:

$$\begin{aligned} \max_{\{y_{j,t}(m)\}} \quad & P_{j,t}Y_{j,t} - \int_0^1 p_{j,t}(m)y_{j,t}(m)dm \\ \text{s.t equation (OA.1).} \end{aligned}$$

The above implies that the price index in layer j is given by $P_{j,t} = \left[\int_0^1 p_{j,t}(m)^{1-\mu_j} dm \right]^{\frac{1}{1-\mu_j}}$.

The demand schedule for each intermediate good producer in layer j of variety m is given by $\left[\frac{p_{j,t}(m)}{P_{j,t}} \right]^{-\mu_j} Y_{j,t}$.

The aggregator of layer $j \in \{1..N\}$ supplies capital to intermediate good producers in layer $j - 1$. The aggregator at layer 0 sells its goods to the household for final consumption.

Intermediate good producers. The intermediate good producer in each layer j of variety m faces the same production technology and conditions as described in Section 3.1. It owns its capital stock, $k_{j,t}(m)$, which depreciates at rate δ , and it hires labor from the household. Now, however, it has an additional degree of freedom: the ability to optimally select its output price. Specifically, the period dividend of an intermediate good producer of variety m in layer $j \in \{0, \dots, N - 1\}$, $d_{j,t}$, is given by:

$$d_{j,t}(m) = p_{j,t}(m) \left[\frac{p_{j,t}(m)}{P_{j,t}} \right]^{-\mu_j} Y_{j,t} - P_{j+1,t} \Phi(i_{j,t}(m)) k_{j,t}(m) - W_t \cdot n_{j,t}(m), \quad (\text{OA.2})$$

where, as before, W_t denotes the real wage per unit of labor. Given that the top layer intermediate good producer's capital is fixed as before ($k_{N,t}(m) = k_{N,0}(m)$), their dividend is similarly given by $d_{N,t} = p_{N,t}(m) \left[\frac{p_{N,t}(m)}{P_{N,t}} \right]^{-\mu_N} Y_{N,t} - W_t n_{N,t}(m)$. Each intermediate good producer chooses its output price, optimal hiring, and investment (except producers at the top firm), to maximize its market value, taking as given wages W_t , its layer price index $P_{j,t}$, $j \in \{0, \dots, N\}$, and the stochastic discount factor of the household $M_{t,t+1}$. Specifically, layer- j firm maximizes:

$$V_{j,t}(m) = \max_{\{n_{j,s}(m) \text{ and } p_{j,s}(m) \text{ iff } j \in \{0, \dots, N\}; k_{j,s+1}(m) \text{ iff } j \in \{0, \dots, N - 1\}\}} E_t \sum_{s=t}^{\infty} M_{t,s} d_{j,s}(m); \quad (\text{OA.3})$$

$$\text{s.t.} \quad (\text{OA.4})$$

$$\left[\frac{p_{j,t}(m)}{P_{j,t}} \right]^{-\mu_j} Y_{j,t} \leq Z_t k_{j,t}(m)^\alpha n_{j,t}(m)^{1-\alpha} \quad (\text{OA.5})$$

$$k_{j,t+1}(m) = (1 - \delta + i_{j,t}(m)) k_{j,t}(m) \text{ if } j \in \{0, \dots, N - 1\}. \quad (\text{OA.6})$$

Market clearing and equilibrium. Compared to Section 3.3, the market clearing conditions of the labor markets, the capital goods markets, and the consumption good market are modified

as follows:

$$\begin{aligned}\sum_{i=0}^N \int_0^1 n_{j,t}(m) dm &= 1, \\ \int_0^1 \Phi(i_{j,t}(m)) k_{j,t}(m) dm &= Y_{j+1,t} \quad \forall j \in \{0, \dots, N-1\}, \\ C_t &= Y_{0,t}.\end{aligned}$$

All other market clearing conditions remain the same. Equilibrium consists of prices, labor, and capital allocations such that (i) taking prices and wages as given, the household's allocation solves (9), and firms' allocations solve (OA.3); (ii) all markets clear; (iii) we are interested in a symmetric equilibrium in which $k_{j,t}(m) = k_{j,t}$, $n_{j,t}(m) = n_{j,t}$, and $p_{j,t}(m) = p_{j,t}$, for all $m \in [0, 1]$.

We calibrate the augmented model at an annual frequency using a calibration identical to that described in Section 5.1. We further impose that $\mu_j = \mu$, $\forall j \in \{0, \dots, N\}$. We are interested in varying the markup parameter μ , and considering the quantitative implications that it has on the spread.

OA.3 A Measure of Supply Chain Competition

To derive a measure of competition that captures not only a firm's own competition environment, but also the competition faced by its direct and indirect suppliers, we combine the information about the supplier-customer relations with the information about the number of competitors each firm has. The FactSet Revere relationships dataset allows us to identify each firm's competitors at any point in time, reported either by the firm itself or by its competitors. We assume that competition relationships are undirected links, meaning that it is sufficient for only one firm to report the relationship. We observe 271,586 competition links in the database. We eliminate links that last less than 90 days and combine relationships where there is a gap of less than 90 days, with a gap defined as the number of days between the end of the previous relationship and the beginning of a new relationship between the same pair of firms.

Let $\mathbf{C}_t$ be an n by 1 column vector that measures the number of each firm's competitors in month t . While it could be a measure of competition, it does not account for competition at the supply chain level. We define our supply chain competition measure as:

$$\hat{\mathbf{C}}_t = \mathbf{C}_t + \sum_{j=1}^J \lambda^j \bar{\mathbf{S}}_t^j \mathbf{C}_t, \quad (\text{OA.7})$$

where $\hat{\mathbf{C}}_t$ is an n by 1 column vector that measures each firm's supply chain competition, $\bar{\mathbf{S}}_t$ is

a customer-supplier adjacency matrix normalized by the number of suppliers that each customer has. In other words, if $\mathbf{S}_t$ is an adjacency matrix of zeros and ones, such that $\mathbf{S}_t(i, j) = 1$ if firm j is a supplier to firm i and $\mathbf{S}_t(i, j) = 0$ otherwise, then $\bar{\mathbf{S}}_t(i, j) = \mathbf{S}_t(i, j) / \sum_{k=1}^{n_t} \mathbf{S}_t(i, k)$. The $\lambda < 1$ parameter discounts the importance of the competition faced by a firm's direct suppliers relative to the competition faced by the firm itself, and the importance of competition faced by indirect suppliers relative to the competition faced by direct suppliers.²

When $J = 1$, the supply chain competition measure $\hat{\mathbf{C}}_{it}$ for a given firm i includes the firm's own number of competitors and the average number of its direct suppliers' competitors. For $J = 2$, the average number of competitors of the firm's suppliers' suppliers is also included in the measure. In our benchmark specification we use $J = 5$, accounting for the competitiveness of indirect suppliers five layers above the firm itself.

OA.4 Cyclicity of Input Prices Using BLS FD-ID Database

Theorem 2 in Section 4 and Table 7 suggest that in the model, the relative prices of (capital) inputs are more cyclical for upstream firms than for downstream firms. Unfortunately, we cannot observe the relative input prices for firms in different production layers. This is because our benchmark analysis accounts for within-sector supply chain, and firm-level input prices data are not publicly available. Nonetheless, we show in Section 7.1.1 that the inter-industry component of the TMB spread is positive and significant, using input/output tables. Therefore, we can test the model prediction related to input prices at the industry level, using relevant data from BLS.

As part of its Final Demand-Intermediate Demand (FD-ID) system, the BLS publishes every month price indices for inputs used by intermediate good producers. The intermediate demand input-price indices are organized in a stage-based system. The system is comprised of four intermediate production stages. To construct this system, the BLS assigns industries into production stages, in a manner that maximizes downstream production flow, while minimizing upstream production. The price index for each stage tracks the prices for various inputs that are mostly commonly used by industries that belong to the intermediate production stage.

The BLS follows a different indexing system than our convention. In the FD-ID stage system, stage 1 refers to the most upstream industries, while stage 4 is the most downstream industries.

²We set $\lambda = 0.9$ in our calculation of the supply chain competition measure. The results are robust to altering this parametrization.

Industries belonging to the fourth stage primarily produce output consumed as final demand (primarily business equipment demand). These industries include, for example, manufacturers of utility vehicles, automobiles, and pharmaceuticals. Industries in the third stage produce output consumed by stage 4 industries. These include motor vehicle parts and semiconductors. Industries assigned to the second stage produce output consumed by stage 3 industries. These include petroleum refineries, natural gas distribution, electricity generation, and manufacturers of plastic materials. Industries assigned to the first stage produce output primarily consumed by stage 2 industries. These include oil and gas extraction, paper mills, and grain farming. More information is available at: <https://www.bls.gov/opub/mlr/2011/02/art1full.pdf>.

There are two limitations in using the FD-ID stage-input prices to test our model. First, the sample is very short. The BLS data is only available starting at 2010-Q1. Second, while the stages of production constructed by the BLS have resemblance to our industry-level upstreamness score, the classifications are also different in two ways: (a) our methodology to assign industries into production layers follows Antràs et al. (2012), and accounts for all possible *weighted* paths from a given industry to the bottom; (b) the bottom layer is defined differently. In order to parallel the economy with a linear supply chain that ends with the household, we define the bottom layer as consumption good producers. The BLS defines the bottom layer as any producer of final demand, which includes (completed) commodities sold not only for personal consumption, but also for capital investment, government purchases, and exports.

Despite the limitations above, we test the model's prediction using the following regression:

$$\Delta p_{j,t} = \text{const} + \beta_{\text{prod}} \Delta \text{Prod}_t + \text{error}, \quad (\text{OA.8})$$

where $\Delta p_{j,t}$ is the quarterly log growth in the input price for stage $j \in \{1..4\}$ (series WPUID51-4), and ΔProd is the quarterly log growth in aggregate productivity. Since the input price indices are taken from BLS, we use (annualized) labor productivity to calculate ΔProd , which is also computed by the BLS.

The results are reported in Table OA.2. With the exception of Stage 2, the input-price productivity beta increases from Stage 4 (the most downstream) to Stage 1 (the most upstream). The productivity beta difference between Stage 1 and Stage 4 is 0.37, and is significant at the 5% level. This suggests that consistent with our model's prediction, input prices are more procyclical for upstream firms.

Table OA.2: **Input-Price Cyclicity for Intermediate Demand Stages**

	Stage 4 (most downstream)	Stage 3	Stage 2	Stage 1 (most upstream)	Diff (1-4)
β_{prod}	0.11	0.25	0.11	0.48	0.37
	(1.64)	(1.64)	(0.45)	(2.09)	(2.08)

The table shows results of the projection $\Delta p_{j,t} = const + \beta_{prod} \Delta Prod_t + error$ where $p_{j,t}$ is the input price index for intermediate stage j , and $Prod$ is annualized labor productivity. Both time-series are quarterly from the BLS. Diff(1-4) refers to the difference in β_{prod} between stage 1 and stage 4. HAC t-statistics are in parentheses. The sample period is from 2010Q1 to 2017Q4.

OA.5 Model Assumptions and Extensions

OA.5.1 Goods transmitted in supply-chains.

In the model, we assume that suppliers sell capital goods to customer firms. Our interpretation of capital in this economy is loose: it is possible to interpret capital goods as machines or as intermediate inputs.³ One can view intermediate inputs as a form of “capital” that depreciates faster (see e.g. Belo and Lin (2011); Jones and Tuzel (2013)). For parsimony, we refrain from introducing to the model two types of capital goods (fixed capital and inventory). Instead, we model production using one form of capital, but we verify that the conditions needed for the mechanism to be applied for inventory inputs are held. There are two crucial ingredients for vertical creative destruction to apply for a given input: (1) the depreciation rate for the good transmitted is less than one, and (2) the good is stored in a stock by the firm (either physically or through a forward contract). These conditions naturally apply for fixed assets. We also confirm that they apply for intermediate inputs. First, we verify in Section OA.1 that if we assume that only intermediate inputs are transmitted, and consequently, alter the calibrated depreciation rate to match inventory carrying costs (i.e., inventory’s depreciation as in Jones and Tuzel (2013)), then the quantitative effect on our results is minor (the TMB spread is 12% per annum under a higher depreciation versus 12.5% under the benchmark calibration). Second, we confirm in the data that on average all layers have positive inventories, and that the ratio of inventory to firm value is the largest for bottom-layer firms. Consistent with the fact that the mechanism can be equally applied to stocks of inventories, we show in Section 6.3 that the TMB spread is larger for firms that carry a higher amount of inventory.

³For example, magnets producer purchases a large stock of rare earths and uses them in production of magnets over a period of one year or more. From our mechanism’s perspective, it is not conceptually different from purchasing a grinding machine, which is used to smooth magnets. Both rare earths and grinding equipment are held in stocks, and depreciate gradually over time.

OA.5.2 Ex-ante Layer Heterogeneity

In the baseline model we introduce as little heterogeneity between the layers as possible, to ensure that any differences in returns arise from network effects and not from any hardwired ex-ante heterogeneity. As previously discussed, many supply chains are very long and extend almost exclusively within sectors. Hence, the TMB spread is not driven by any specific industry. Nonetheless, top layer firms could exhibit a greater amount of durability in their capital input and/or output because of the nature of their technology. We modify the benchmark model parameters such that firms with higher vertical position have either higher durability due to lower depreciation, or higher degree of adjustment costs. The results are reported in Table OA.3. When the depreciation rate (capital adjustment costs) of the top layer is *twice* as low (high) as that of the bottom layer, the marginal increase in the TMB spread compared to the baseline model is quantitatively very small, about 1%. Thus, we do not consider ex-ante heterogeneity as a *quantitatively* important source for the TMB spread.

Table OA.3: TMB Spread and Ex-Ante Layer Heterogeneity

	(1) Benchmark Calibration	(2) δ_{mid} fixed $\delta_0 > \delta_{N-1}$	(3) ϕ_{mid} fixed $\phi_0 < \phi_{N-1}$
Mean Excess Returns (%)			
Layer 5	16.07	16.83	16.41
Layer 4	12.01	11.35	13.16
Layer 3	9.42	8.97	9.78
Layer 2	7.15	6.97	6.87
Layer 1	5.23	5.23	4.69
Layer 0	3.59	3.66	2.89
Spread (5-0)	12.49	13.17	13.51

This table reports how the benchmark results would change if we allow for ex-ante heterogeneity between the layers of production. It shows the model-implied excess returns for the different production layers under three calibrations: (1) The benchmark calibration, in which the depreciation rate and adjustment costs are homogeneous for all layers (i.e., $\delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta$; $\phi_0 = \phi_1 = \phi_2 = \phi_3 = \phi_4 = \phi$), (2) A calibration with a mean-preserving spread of depreciation rates, in which the top layers' capital depreciates slower (i.e., $\delta_0 = 1.5\delta, \delta_1 = 1.25\delta, \delta_2 = \delta, \delta_3 = 0.75\delta, \delta_4 = 0.5\delta$), (3) A calibration with mean-preserving spread of capital adjustment costs, in which the top layers incur larger adjustment costs (i.e., $\phi_0 = 0.5\phi, \phi_1 = 0.75\phi, \phi_2 = \phi, \phi_3 = 1.25\phi, \phi_4 = 1.5\phi$).

OA.5.3 Alternative Network Specifications

To focus on the vertical dimension of the production network and to keep the model parsimonious, we use a perfectly vertical supply chain. Our decision to have a perfectly vertical supply chain, where the capital of the top layer is fixed, stems from two reasons. First, this assumption

allows us to capture a salient feature of the data, in which the spread between the average return of layer five and four is relatively large. As the top layer cannot reduce the cyclical of its dividends via investment, it bears an extra premium for that. Second, this modeling assumption is also consistent with empirical evidence that the labor share of output is higher for investment good producers (see e.g. Basu et al. (2012)). In this section, we consider two alternative network specifications: a pyramidal network, and a chain which includes a loop at the top (i.e., the capital of the top layer is not fixed). When the elasticity of substitution between the consumption good is close to unity, the pyramidal structure yields results which are quantitatively close to the benchmark chain model. We also confirm that a positive and sizable model-implied TMB spread exists when the chain has a loop at the top (that is, the top layer firm sells capital to itself, and to the layer below it, making the capital of the top layer time-varying).

Pyramidal Structure. A salient feature of the data, demonstrated in Table 14 is that the production network features a pyramidal shape, in line with the notion of the Hayerkian Triangle. More firms populate downstream layers. To capture the main economic force, and for tractability reasons, our benchmark model features a strict chain where each layer is populated by a single representative firm. In light of the empirical regularity, we wish to qualitatively assess the impact of the pyramidal structure on the TMB spread. To that end, we solve a simple model that features only two layers of production. The top layer is populated by a single firm, while the bottom layer is populated by two final consumption good producers. Denote the two final consumption goods by c_1 and c_2 , respectively, the period utility of the bundle (c_1, c_2) is given by the CES function:

$$\left[(1 - \eta)c_{1t}^{1-\frac{1}{\rho}} + \eta c_{2t}^{1-\frac{1}{\rho}} \right]^{\frac{1}{1-\frac{1}{\rho}}},$$

where the parameter ρ captures the degree of substitutability between the consumption goods. Specifically, for $\rho = 1$ the period utility reduces to the Cobb-Douglas function. When $\rho = 0$, the function reduces to a Leontief function, and the goods are perfect complements. When ρ approaches infinity, the function becomes linear, and the goods are perfect substitutes. We set $\eta = 0.5$, so that both consumption goods receive an equal weight. The CES period utility function is nested inside an Epstein-Zin preferences over today's consumption and the continuation utility. The output of the top layer is sold to both bottom-layer firms:

$$Y_{top,t} = \Phi(i_{c1,t})k_{c1,t} + \Phi(i_{c2,t})k_{c2,t}.$$

This structure captures a very simple triangle-shaped economy.

Table OA.4: TMB Spread implied From A Pyramid versus Chain Network

ρ	σ	TMB-Pyramid	TMB-Chain
1.2	0	7.56	7.56
1.2	0.15	6.09	6.03
1.2	0.30	4.45	4.34
0.8	0	7.56	7.56
0.8	0.15	5.87	6.03
0.8	0.30	3.95	4.34

This table compares the TMB spread in a model with a production chain, like in the benchmark case, and the spread in a pyramidal model, in which a supplier has two customers. The specification of the chain model is identical to the benchmark model, with the modification of setting N to 2, and an introduction of a firm specific shock. In both the pyramidal and chain models, the production function for any firm i is given by $Y_{i,t} = Z_{i,t} k_{i,t}^\alpha n_{i,t}^{1-\alpha}$, where $Z_{it}/Z_{i,t-1} = \underbrace{Z_t/Z_{t-1}}_{\text{common trend}} \cdot \underbrace{\exp(\sigma_{idio}\varepsilon_{i,t})}_{\text{layer-specific shock}}$, and $\varepsilon_{i,t}$ is i.i.d. standard Gaussian firm-specific shock. In the pyramidal

model, the TMB spread is defined as the difference between the return of the top-layer firm and the value-weighted return of the two bottom-layer firms. The table shows the results when the CES parameter ρ takes the values $\{0.8, 1.2\}$, and the firm-specific volatility parameter σ_{idio} takes the values $\{0, 0.15, 0.3\}$.

Table OA.4 compares the TMB spread in a strict chain model with the spread in a pyramidal model. The triangle shape has an ambiguous effect on the TMB spread. In general, the higher the substitutability between the consumption goods, the higher the spread. Specifically, we find that if the CES parameter is above (below) unity, the pyramidal structure boosts (shrinks) the TMB spread, relative to a chain structure. Moreover, the triangle structure has no effect on the spread if the period utility between the two consumption goods is Cobb-Douglas (CES parameter equals to one) and/or if the firm specific productivity shocks are sufficiently small. In reality, empirical estimates of the CES parameter for the bottom layer (i.e., complementarity or substitutability between durable and non-durable consumption goods) is not much different from unity (see e.g. Yang (2011)). For instance, Ogaki and Reinhart (1998) find the CES parameter to be around 1.2, whereas Yogo (2006) estimates it at around 0.8. As a result, the triangle shape should have a quantitatively small effect on the TMB spread.

Supply chain with a loop. Our benchmark model features a perfectly vertical supply chain. It also implies that the capital of the top layer is fixed over time because it does not have suppliers. This assumption helps to capture the sizable spread between layer 5 and 4, and is motivated by the fact that top-layer firms are more labor intensive. However, the TMB spread exists even when the top layer firms can purchase inputs from other top layer firms, allowing a loop within layer 5. Specifically, we solve a variation of the benchmark model in which the top layer's capital is time-varying. The output of the top layer is sold as capital input for the layer below it (layer 4),

as well as to itself (layer 5, thus featuring a circle). We confirm that a positive and sizable TMB exists under this circular model.⁴

OA.5.4 Transition between Layers: A Quantitative Assessment

The model in Section 3 features a linear supply chain with $N + 1$ production layers. For tractability, the model assumes a single representative firm in each layer, and that a firm's vertical position is fixed over time. In the data however, the vertical position of a firm is not static. While the vertical position is a persistent measure, it is also dynamic (time-varying). The purpose of this section is to bridge the gap between the model and the data on this matter. We provide comparative statics that quantitatively assess the impact of potential transitions of firms between layers on the model-implied TMB spread. The computation combines both model-implied risk premia and empirical information on transition probabilities. We show that allowing for dynamic transitions reduces the TMB spread, but the effect is quantitatively very small. In fact, given that the model-implied TMB spread is slightly higher than the data, accounting for the probability of transitions brings the model-implied TMB spread even closer to the data.

Let R be an $(N+1) \times 1$ vector of model-implied one-period expected returns. $R(j)$ is the expected return of a firm that belongs to layer $j \in \{0, \dots, N\}$. Assume that firm i currently operates in layer j in period $t - 1$. Right at the beginning of period t , before operation starts, the firm may stay in the same layer or move to a different layer. If the firm remains in layer j , its expected return from t to $t + 1$ is $R(j)$. If on the other hand, right at the beginning of period t the firm transitions to layer $k \neq j$, its expected return from t to $t + 1$ becomes $R(k)$, in a comparative static sense. Let P be an empirically-driven between-layer transition probability matrix. $P(j, k)$ is the probability that a firm which is currently in layer j transitions to layer k next period, and $\sum_{k=0}^N P(j, k) = 1$. The expected return of firm i between t and $t + 1$, accounting for transition dynamics, is given by the multiplication of the transition probability vector for layer j and the vector of expected returns R . In other terms, we can adjust the expected return of a firm that currently belongs to layer j , to account for the possibility of a dynamic transition by:

$$E_t^{adj}[R_{i,t+1}|i \in j] = \sum_{k=0}^N P(j, k)R(k) \quad (\text{OA.9})$$

To implement the equation above we make two preliminary computations. First, we construct

⁴The spread is attenuated to about 4.3% per annum.

the transition matrix P from the monthly data. The matrix of monthly transition probabilities between layer j and layer k , where $j, k \geq 1$ is taken from Table OA.10. Denote this matrix as $P_{1..N,1..N}$. We augment this matrix to accommodate layer zero as well. In our empirical specification, a firm belongs to layer zero if it produces final consumption (staples or discretionary) products (i.e., based on fixed sectoral affiliation). Thus, a firm that belongs to layer zero cannot transition upstream, and a firm that operates in a layer above zero cannot transition to layer zero. Overall $P = [1, 0_{1 \times N}; 0_{N \times 1}, P_{1..N,1..N}]$. Second, we obtain the vector R of monthly expected returns from the calibrated model solution, reported in Table 5. Because the estimated transition probabilities are monthly, we adjust the model-implied annual expected returns to monthly terms. If $R^a(j)$ is the annual model-implied expected return for a firm in layer j , then $R(j) = (1 + R^a(j)/100)^{\frac{1}{12}}$. We then apply equation (OA.9) for all layers.

Table OA.5: **Model-Implied Expected Return Adjusted for Transition Dynamics**

	Model (unadjusted)	Model (adjusted)
Layer 5	1.250	1.165
Layer 4	0.950	0.929
Layer 3	0.753	0.745
Layer 2	0.577	0.577
Layer 1	0.426	0.427
Layer 0	0.294	0.294
TMB (monthly)	0.955	0.870
TMB (annual)	12.49	11.31

The table shows the model-implied monthly expected returns of different layers before and after adjusting for transition dynamics. The unadjusted monthly expected returns are computed from layer j using Table 5 as $(1 + R^a(j)/100)^{\frac{1}{12}}$. The adjusted monthly expected return is computed from Equation OA.9. The monthly TMB is computed as the difference between layers 5 and 0. The annual TMB is computed as the difference between the annualized returns for layers 5 and 0: $(1 + E[R_5]/100)^{12} - (1 + E[R_0]/100)^{12}$, where $E[R_j]$ is a monthly (adjusted) expected return.

Table OA.5 shows the raw and adjusted model-implied expected returns for all layers. For layers 5, 4, and 3 the adjusted expected return is slightly smaller, because the probability of transitioning upstream is smaller than transitioning downstream, and downstream firms have lower risk premia. For layer 1, and 2 the opposite is true (for layer 2 the effect is too small to be visible in the table). For these downstream firms, the probability of moving upstream is slightly larger, thereby increasing the adjusted expected return. The expected return of layer 0 is unchanged because layer 0 firms do not transition. We then compute the annual TMB as the spread between the annualized

expected returns (unadjusted or adjusted) of layers 5 and 0. The adjusted TMB spread decreases by a small amount of about 1.2% to 11.31%. The spread falls because the adjusted return of layer 5 must decrease (the firm cannot move upstream), while the adjusted return of layer 0 is fixed (the firm remains in layer 0). Table 5 shows that the empirical TMB spread (annualized) is 11.27%. Thus, the adjusted model-implied spread fits precisely the point estimate in the data.

OA.6 Vertical Position Computation Using Input-Output Tables

In this subsection, we explain the advantages and disadvantages of using BEA Input-Output tables relative to the FactSet Revere firm-level data about supplier-customer relationships that is used in the paper. We also explain the methodology for constructing the alternative vertical position measure using the Input-Output tables.

FactSet Revere Relationships Database vs. BEA Input-Output Tables The advantage of the I-O tables is that they span a longer period than the FactSet Revere database. They also contain information about the dollar flow of goods between industries. As a result, the vertical positions computed in this subsection alleviate two concerns about our baseline results: (1) they account for the strength of the links between any two industries, not only for its existence; and (2) they account for a weighted average of all possible “paths” from one industry to consumers, as opposed to the minimum path as in the baseline results. However, the I-O tables also have serious limitations. Most importantly, they capture only the across-industry supply chains, but not the within-industry supply chains. Many production chains extend almost exclusively within sectors. Consequently, firms can have a high vertical position even if their industry (as a whole) does not. Since the TMB spread estimated using the I-O tables cannot capture any intra-industry component of the spread, our conjecture is that it should be smaller compared to the spread estimated using firm-level data. Another limitation of the I-O tables is that the benchmark tables, which classify industries/commodities with the granularity required for our study, are published by the BEA only once every 5 years. For example, benchmark Make and Use tables are available for 2012, 2007, 2002, 1997, and so on.

Methodology. We compute vertical positions from Make and Use tables using two steps, as described next. Let N_t be the number of industries considered by the BEA analysis at year t , and let C_t be the number of commodities considered. The dimension of the Make table, denoted by M_t , is $N_t \times C_t$. The dimension of the Use table, denoted by U_t , are $C_t \times (N_t + 1)$. The first N_t

columns of the Use table capture the dollar flow of a commodity into each industry (as an input), whereas the last column shows the flow of the goods for consumption.

The first step to compute industry vertical positions at time t , is to normalize the Use and Make tables such that the sum of elements in each row is one, as follows:

$$\begin{aligned}\tilde{M}_{j,i,t} &= \frac{M_{j,i,t}}{\sum_{k=1}^{C_t} M_{j,k,t}}, \quad \forall i \in \{1, \dots, N_t\}, j \in \{1..C_t\}, \\ \tilde{U}_{i,j,t} &= \frac{U_{i,j,t}}{\sum_{k=1}^{N_t+1} U_{i,k,t}}, \quad \forall i \in \{1, \dots, N_t\}, j \in \{1..C_t\}.\end{aligned}$$

The element (j, i) in $\tilde{M}_t$ captures the share of commodity i produced by industry j . Likewise, the element (i, j) in $\tilde{U}_t$ captures the share of commodity i used by industry j . That share takes into account that some amount of the commodity flow may also be used for final consumption.

Let $\tilde{U}_{C_t \times N_t}$ be the first N_t columns of $\tilde{U}$. If entire row i in this sub-matrix is zero, then commodity i is used strictly for consumption. Any industry that produces solely this commodity should have the lowest vertical position (i.e., it belongs to the bottom-most layer of production). If the sum of elements in row i in matrix $\tilde{U}_{C_t \times N_t}$ is positive, then the commodity is used for intermediate stages of production by other industries. Thus, any industry producing this commodity, perhaps along with other commodities, should have a vertical position above the lowest score.

The second step involves combining the information from the normalized Make and Use tables to form upstreamness scores. Following Antràs et al. (2012), we compute upstreamness as follows:

$$\tilde{V}P_t^{Ind} = (I_{N_t \times N_t} - \tilde{M}_{N_t \times C_t} \cdot \tilde{U}_{C_t \times N_t})^{-1} \iota, \quad (\text{OA.10})$$

where $\tilde{V}P_t^{Ind}$ is a $N_t \times 1$ vector of industries' upstreamness scores at time t , I is the identity matrix, and ι is an $N_t \times 1$ vector of ones. The minimal value in vector $\tilde{V}P_t^{Ind}$ is 1, but the maximal value is unbounded. As we are using Make and Use tables from different years, there is a significant variation in the maximum values. For portfolio formation, and for meaningful comparison between years, it is important to keep the vertical positions score of different years on the same scale. Thus, we define industries' vertical position, VP_t^{Ind} at time t , as a linear transformation of the upstreamness score $\tilde{V}P_t^{Ind}$:

$$VP_t^{Ind} = A \cdot \frac{\tilde{V}P_t^{Ind} - \min(\tilde{V}P_t^{Ind})}{\max(\tilde{V}P_t^{Ind}) - \min(\tilde{V}P_t^{Ind})}, \quad (\text{OA.11})$$

where VP_t^{Ind} is an $N_t \times 1$ vector of (normalized) vertical positions. The transformation ensures that the minimal vertical position score is 0 and the maximal one is a scalar A . For comparison with our baseline results we set $A = 5$.

Further Details. We use the Make and Use Tables for years 2012, 2007, 2002, 1997, 1992, 1987, 1982, and 1977. For each year, we remove in the Use and Make tables industries that relate to the local, state and federal Government (S001–S007), as well as Other Services (NAICS 81). The final consumption usage of each commodity (i.e., column $N_t + 1$ in the Use table) is taken from Personal Consumption Expenditures column in the BEA Use files (F01000 or 910000). In case that the sum of a particular row in the Use matrix U_t is zero, we use a row of zeros in the respective use share table $\tilde{U}_t$. When computing the use share table $\tilde{U}_t$, if a particular industry supplies more than 90% of its output to itself (i.e. $\tilde{U}_t(i, i) > 0.9$), we exclude it from the table. We apply this filter because such industries are almost completely disconnected from the rest of the economy (i.e., close to an autarky industry). Including such industries in the vertical position computation not only generates an extreme vertical position score for the industry itself, but also artificially elevates the vertical position for other industries that supply to it, directly or indirectly. As part of the vertical position scores, we multiply the use and make share matrices (see Equation OA.10). Denote the $N_t \times N_t$ multiplication matrix by Δ_t (i.e. $\Delta_t = \tilde{M}_{N_t \times C_t} \cdot \tilde{U}_{C_t \times N_t}$). For each row i in this Δ_t matrix, if a certain element indexed by (i, j) is less than 1% (i.e., the amount transferred from industry i to industry j is negligible) we set it to zero, but then reallocate that small amount into all other non-zero entries of row i . This allows us to filter noise in the input/output tables and to focus the analysis on economically meaningful inter-industry flows. We use the data provided by the BEA to map between I-O codes of industries into NAICS codes (for 1997, 2002, 2007, and 2012 tables), and to map between I-O codes of industries into SIC codes (for 1977-1987 tables). For years 1997-2017, when assigning vertical positions for firms, we first try to match a firm's 6-digit NAICS code to a 6-digit NAICS industry for which vertical position is available (the success of this matching depends on the granularity of the input/output tables). If no such vertical position was found, we try to match the firm's 5-digit NAICS code to an available industry. If no match was found we turn to the 4-digit code, and so on, up to the 2-digit code. Similarly, for the years 1977-1987 we first try to match firms to 4-digit SIC code industries, and if no match was found we use the 3-digit code or the 2-digit code, in that order. We consider only stocks with share codes 10, 11 or 12.

OA.7 Vertical Position Computation Compustat Segment Data

In this subsection, we explain the advantages and disadvantages of Compustat Segment Data relative to the FactSet Revere firm-level data about supplier-customer relationships that is used in the paper. We also explain the methodology for constructing the alternative vertical position measure using the Compustat Segment Data.

FactSet Revere Relationships Database vs. Compustat Segment Data There are two advantages to using the Compustat Segment data. First, the data starts at early 1980s. This allows to construct a variation of the TMB spread for a relatively long sample period of over 30 years. Second, the data includes the flow of sales for each supplier-customer relationship. This allows to construct a vertical position measure that accounts for the strength of the link between any two firms.

However, the construction of vertical positions using Compustat Segment data has significant disadvantages. Because only critical customers are reported, this truncation of the links suggests that there are far fewer relationship in the database. For comparison, the FactSet Revere database has about 24,000 supplier-customer relationships covering 4,000 firms in 2003, the year when the database started, while the Compustat segment data has only 2,837 relationships relationships covering 847 customer firms in 2003. Because of the missing links in the Compustat Segment data, the observed supply chains are significantly shorter, making it harder to measure a firm's true vertical position, and harder to identify sufficiently many firms with a high vertical position score.

Methodology. We compute vertical positions of firms from adjacency matrices obtained from Compustat Segment data, as described next. Let N_t be th number of firms considered by Compustat Segment at year t . Let A_t be an adjacency matrix of size $N_t \times N_t$. Cell A_t^{ij} show the dollar sales from firm i to firm j in year t .

Let Z_t be a column vector of zero and ones of size N_t . We set Z_t^i to 1 if firm i belongs to layer zero, that is, belongs to the sector of consumer discretionary (GICS code 25) or consumer stables (GICS code 30), or 0 otherwise. To be consistent with the benchmark analysis, we wish to ensure that if $Z_t^i = 1$ then firm i belongs strictly to layer zero. This requires adjusting matrix A_t , such that if $Z_t^i = 1$, then row i in matrix A_t is set to all zeros. Denote by $\hat{A}_t$ a matrix of dimensions $N_t \times (N_t + 1)$, constructed as $\hat{A}_t$ with vector Z_t appended from the left. In other terms, $\hat{A}_t = [A_t, Z_t]$. Next, we normalize matrix $\hat{A}_t$ such that the sum of elements in each row is one, as

follows:

$$\tilde{A}_{i,j,t} = \frac{\hat{A}_{i,j,t}}{\sum_{k=1}^{N_t+1} \hat{A}_{i,k,t}}, \quad \forall i, j \in \{1, \dots, N_t\}.$$

The element (i, j) in $\tilde{A}_t$ capture the share of sales of firm i to firm j , taking into account that the firm may sell directly to consumers ($Z_t^i = 1$), in which case $\tilde{A}_{i,N_t+1,t} = 1$.

Let $\tilde{A}_{N_t \times N_t}$ be the first N_t columns of $\tilde{A}$. If the sum of elements in row i of $\tilde{A}_t$ is positive (and equals to one), then the firm has vertical position of one or above (i.e., it is a direct or indirect supplier of a layer zero firm). If the sum of elements in row i is zero, the firm is a consumption-good producer belonging to layer zero. To compute vertical position scores that take into account the weight of suppliers' sales to each of its customers, we follow Antràs et al. (2012), similar to our analysis of BEA Input-Output Tables. Let the vertical position scores be:

$$\tilde{V}P_t^i = (I_{N_t \times N_t} - \tilde{A}_{N_t \times N_t})^{-1} \iota, \quad (\text{OA.12})$$

where $\tilde{V}P_t^i$ is an $N_t \times 1$ vector of firms' vertical position scores at time t , I is the identity matrix, and ι is an $N_t \times 1$ vector of ones. Since the maximal vertical position is unbounded, and oscillates between different years, we normalize the vector $\tilde{V}P_t^i$. Specifically, we apply the same linear transformation of Equation (OA.11) onto $\tilde{V}P_t^i$, and define firms' vertical position as:

$$VP_t^i = A \cdot \frac{\tilde{V}P_t^i - \min(\tilde{V}P_t^i)}{\max(\tilde{V}P_t^i) - \min(\tilde{V}P_t^i)}. \quad (\text{OA.13})$$

The vector VP_t^i is an $N_t \times 1$ vector of normalized vertical positions. To be consistent with the BEA Input/Output tables analysis, and also with our baseline results, we set A to 5.

Further Details. The Compustat Segment database reports only the most important customers of each firm at an annual frequency. We filter the firms that enter the adjacency matrix A_t every year in a similar fashion to the filtering applied in our benchmark analysis. Specifically, we remove financial firms (GICS = 40), real estate (GICS = 60), as well as conglomerates. To ensure the stability of links, we apply a filter following Fama and French (1992), which requires that the customer firm exists in the CRSP sample for at least 24 months prior to the inclusion in the adjacency matrices. The number of firms that comprise the adjacency matrix (i.e., for which customer data is available) is too small for vertical position computations in the early 80s. Between 1984 to 1985 the number of firms covered jumps by about 200 firms. Thus, we start the computation of vertical positions in 1985.

OA.8 TMB: Is it a Priced Risk Factor?

We demonstrate in this section that the TMB spread is a novel empirical finding: it is unexplained by existing spreads/factors. Nonetheless, it does not emerge as a novel determinant (factor) for stock prices.

We start by considering the TMB alphas implied from five well-known factor models: the classic CAPM, the Fama and French (1993b) three-factor model, the Carhart (1997) four-factor model, the Hou, Xue, and Zhang (2015) q -factor model, and the Kogan and Papanikolaou (2014) IMC factor. The results are reported in the first five columns of Table OA.6. The alpha of the spread is mostly not affected by controlling for the factors, and it is still positive and statistically significant at least at the 10% significance level.⁵ Qualitatively, the TMB portfolio has positive betas with respect to the market excess return, and to the HML, SMB, and MOM factors. It has a significantly negative loading on the investment factor of the q -factor model (I/A). The loading on the investment-minus-consumption (IMC) factor of Kogan and Papanikolaou (2014) is significantly positive, which is to be expected because the bottom layer consists of consumption-good producers. However, the alpha of the TMB portfolio is still economically and statistically significant, suggesting that the TMB spread is distinct from the IMC factor. We stress that the TMB spread is mostly within the investment sector, instead of a spread between the investment and consumption sectors. In all, the TMB spread cannot be explained by usual unconditional factor models.

In addition, we examine whether our TMB spread simply reflects the durable-nondurable spread (DMN) studied by Gomes, Kogan, and Yogo (2009). The last column of Table OA.6 shows the result of a univariate regression, where the TMB spread is regressed on the DMN. It reveals that these two spreads are largely orthogonal to each other.⁶

In all, table OA.6 demonstrates that the TMB spread cannot be explained by well-established unconditional asset-pricing models. For all unconditional factor model specifications, the TMB

⁵All alphas are significant at the 1% level if we do annual rebalancing (see Table OA.17 in the Online Appendix). If we use equal-weighted portfolios, most alphas are significant at the 1% level (see Table OA.15 in the Online Appendix).

⁶Both the IMC factor (Kogan and Papanikolaou (2014)) and the DMN spread (Gomes, Kogan, and Yogo (2009)) are based on industry classification derived from the benchmark Input-Output tables compiled by the Bureau of Economic Analysis. We construct sector portfolios (consumption vs. investment, durable vs. nondurable & service) using the industry classification table developed by Gomes, Kogan, and Yogo (2009), and use the value-weighted returns of these portfolios to estimate the time series of IMC and DMN. We thank Motohiro Yogo for making this table available. We also thank Ken French and Chen Xue for making their factor data available.

Table OA.6: Time series regressions of the TMB portfolio return

	(1)	(2)	(3)	(4)	(5)	(6)
Rm-Rf	0.078 (0.68)	0.016 (0.11)	0.097 (0.69)	0.007 (0.05)	-0.280* (-1.98)	
SMB		0.261 (1.02)	0.211 (0.84)	0.348 (1.37)		
HML		0.008 (0.03)	0.119 (0.52)			
MOM			0.227** (2.04)			
I/A				-0.917** (-2.56)		
ROE				-0.069 (-0.25)		
IMC					0.784*** (3.94)	
DMN						0.042 (0.49)
Constant	1.010* (1.98)	0.993* (1.93)	0.952* (1.88)	1.100** (2.15)	1.167** (2.42)	1.058** (2.08)
R^2	0.004	0.014	0.051	0.072	0.128	0.002

This table presents the results of the time series regressions of the value-weighted TMB spread on various risk factors and the durable-nondurable spread. The TMB portfolio is constructed by longing the value-weighted portfolio of companies with a vertical position of five or above, and shorting a value-weighted portfolio of companies with a vertical position of zero. The returns are computed from November 2003 to February 2013. Rm-Rf is the excess return of the market portfolio. SMB and HMB are the size (small-minus-big) and book-to-market (high-minus-low) factors in the Fama and French (1993b) three-factor model, and MOM is the momentum factor in the Carhart (1997) four-factor model. I/A and ROE are the investment and profitability factors, respectively, in the q -factor model of Hou, Xue, and Zhang (2015). IMC is the investment-minus-consumption factor in the Kogan and Papanikolaou (2014) two-factor model. DMN is the return difference between the portfolio of durable goods producers and the portfolio of nondurable goods and service producers. Both IMC and DMN are constructed based on Gomes, Kogan, and Yogo (2009). T-statistics are reported in parentheses. Significance at the 5% and 10% levels are indicated by ** and * respectively.

alpha is positive and significant. There are three ways to interpret this result.

First, it could be the case that while unconditional specifications of factor models fail to capture the spread, a conditional specification can. This can arise if the TMB has time-varying exposure to the underlying risk captured by the factors. Section 5.2.3 suggests that this holds within the model. We show in Section 5.2.3 that the betas of layers' returns to the underlying shock are not fixed, and the relation between TMB and the underlying risk factor is non-linear. As a result, the model can capture a significant fraction of the CAPM α .

Second, our risk-based explanation for the TMB spread, detailed in Section 3 to Section 5, relies on a single specific factor: aggregate productivity. The evidence in Section 2.2 shows that the spread *can* be explained by exposures to aggregate productivity, which arise endogenously in

our model. Existing factor (e.g., SMB or HML) fail to fully explain the spread as the factors are noisy approximations for the underlying productivity shock(s).

Third, it could be that TMB captures a new source of risk which is a distinct determinant (factor) of stock returns. We examine this possibility in the rest of this section. We check whether the TMB spread is priced in the cross-section of stock returns.

To undertake this analysis, we start by assuming (WLOG) that the log stochastic discount factor m_t is linear in the underlying factors, and takes the following form:

$$m_t = 1 - \lambda_{TMB}TMB_t - \lambda'_f f_t. \quad (\text{OA.14})$$

Here, λ_{TMB} is the parameter of interest that measures the market price of risk of the (potential) TMB factor denoted by TMB_t . f is a $k \times 1$ column vector that contains any additional (known) factors, and λ_f is a $k \times 1$ column vector capturing the respective factor risk premia for f . In our implementation, f is either only the excess market return, or the Fama and French (1993a) three factors. Finally, all factors underlying equation (OA.14) are demeaned.

We estimate the risk factor loadings in equation (OA.14) by generalized method of moments (GMM) using the following moment condition:⁷

$$\mathbb{E} [M_t r_{i,t}^e] = 0, \quad (\text{OA.15})$$

where $r_{i,t}^e$ denotes the excess return of test asset i at time t . We employ two sets of test assets to estimate the factor loadings. First, we estimate the risk factor loadings using 25 value-weighted portfolios double sorted on size and book-to-market. Second, following the suggestion of Lewellen, Nagel, and Shanken (2010), we also estimate the factor loadings using a set of 42 portfolios that augments the first set of test assets with the Fama-French 17 value-weighted industry portfolios. This allows us to break the strong factor structure inherent in the returns of the 25 portfolios sorted on size and book-to-market. We use two versions of the TMB spread. First, we use the benchmark TMB spread obtained from FactSet Revere database. Because the time series is relatively short, we also use the TMB spread implied by the BEA input/output tables. This spans from January 1974 to December 2017. Table OA.7 reports the results of the analysis described above. For brevity, the table displays only the coefficient of interest λ_{TMB} , its t-statistic, as well as the mean absolute pricing error (MAE) from each GMM estimation of equation (OA.15).

⁷The GMM procedure is equivalent, under the appropriate weighting matrix, to a two stage Fama and MacBeth (1973) procedure.

Table OA.7: Market Price of Risk for TMB from GMM Estimation

Panel A: TMB based on FactSet Revere Database:

	25 BM/Size Portfolios		25 BM/Size + 17 Industry Portfolios	
	CAPM + TMB	FF3F + TMB	CAPM + TMB	FF3F + TMB
λ_{TMB}	7.022	9.501	1.994	1.619
$t(\lambda_{TMB})$	[2.90]	[3.67]	[0.94]	[0.73]
MAE	0.175	0.167	0.242	0.240

Panel B: TMB based on BEA Input/Output Tables:

	25 BM/Size Portfolios		25 BM/Size + 17 Industry Portfolios	
	CAPM + TMB	FF3F + TMB	CAPM + TMB	FF3F + TMB
λ_{TMB}	13.808	3.019	5.108	-1.725
$t(\lambda_{TMB})$	[4.76]	[1.07]	[4.10]	[-1.30]
MAE	0.630	0.440	0.854	0.561

The table shows the GMM estimation result of Equation (OA.7). For each cross section and factor model specification f , the table reports the market price of risk of TMB, λ_{TMB} , its statistic, t , and the mean absolute pricing error. In columns labeled by CAPM+TMB, f includes only the excess return of the market portfolio. In columns labeled by FF3F+TMB, f includes the market excess return, SMB and HML. Panel A reports the results of the estimation using the benchmark TMB spread implied from FactSet Revere database over 2003/11–2013/02. Panel B reports the results of the estimation using TMB implied from BEA input/output tables from 01/1974–2017/12.

Panel A of Table OA.7 shows that using the benchmark TMB spread, TMB is priced positively in the cross-section of stock returns. However, this result is only significant using the narrow cross-section of 25 book-to-market by size portfolios. By expanding the cross section to include industry portfolios, the factor risk premium of TMB drops and becomes statistically insignificant. Because the sample period in Panel A is relatively short, we confirm using a longer sample in Panel B that TMB fails to be consistently priced the cross-section of equities. In Panel B, whenever the vector f includes only the market excess return, TMB bears a positive and significant market price of risk, independently of the cross-section used. However, when the vector f is expanded to include SMB and HML, the TMB factor loses significance.

In sum, for some configurations the factor risk premium of TMB is positive and significant. This is broadly consistent with the positive alphas shown in Table OA.6, and with the model-implied CAPM α . Upon a closer inspection, the TMB factor risk premium is subsumed by the size and book-to-market factors. We cannot conclude that TMB represents a new factor. Consistently, the failure of unconditional factor models to capture the spread is replicated in the model, without

assuming multiple aggregate shocks.

OA.9 Ruling Out Mechanical Results

Table 14 shows that the number of firms in layer j largely falls as the vertical position j rises. As a result, the top layer is thinly populated compared to layer 0. This section rules out the possibility that the TMB is positive and large primarily because the number of firms between the top and the bottom layers are very uneven.

We start the discussion by examining the relation between the construction of the vertical position measure, and the number of firms in each layer that result from the construction. We consider the following Monte-Carlo experiment. Fix the number of firms in the economy to $\bar{N}$. Fix the number of firms in layer 0 to N_0 . Let A denote an adjacency matrix of a network. Cell $A(i, j) = 1$ iff firm i supplies to firm j . Let p be a density parameter of the network A . Specifically, p captures the ratio between the actual links in the network, and the potential number of links given by $N \cdot (N - 1)$. The experiment involves M iterations. In iteration m of the experiment, we follow the following steps (experiment 1):

- (1) Randomize the adjacency matrix $A(m)$ such that the density of A is always p . This can be easily implemented by setting cell $A(i, j)$ to 1 iff $\varepsilon_{i,j}(m) < p$ where $\varepsilon_{i,j}(m) \sim U(0, 1)$.
- (2) Randomize layer 0 of iteration m by randomly selecting N_0 firms (out of $\bar{N}$).
- (3) Compute the vertical position of each firm using $A(m)$ and Equation (1).
- (4) Count the number of firms in each layer $j \in \{0..5\}$. Firms in layer 5 have a vertical position of 5 or above.

We set $\bar{N}$ to the average network size per month (the sum of Column N in Table 14). We set N_0 to 1067 (the number of firms producing consumption products, as shown in Table 14). We consider three values of p : $3.61 \cdot 10^{-4}$ (identical to the average density in the data), $1.44 \cdot 10^{-3}$ (four times the empirical density), and $1.81 \cdot 10^{-4}$ (half the empirical density) and. We set the number of iterations M to 100. We report the distribution of firms in each layer across the M Monte-Carlo iterations in Table OA.8.

From Panel A of Table OA.8 shows the results when we use the same density parameter p as in the data. The distribution of firms across the layers features two deviations from the empirical

Table OA.8: Monte-Carlo: Random Networks and Vertical Position Distribution

Number of Firms		
<i>Panel A: $p = 3.61 \cdot 10^{-4}$</i>		
Layer 5	927	[884, 971]
Layer 4	33	[24, 45]
Layer 3	98	[84, 115]
Layer 2	266	[248, 291]
Layer 1	623	[594, 651]
Layer 0	1067	[1067, 1067]
<i>Panel B: $p = 1.44 \cdot 10^{-3}$</i>		
Layer 5	24	[17, 32]
Layer 4	1	[0, 2]
Layer 3	18	[13, 27]
Layer 2	373	[342, 399]
Layer 1	1535	[1507, 1570]
Layer 0	1067	[1067, 1067]
<i>Panel C: $p = 1.81 \cdot 10^{-4}$</i>		
Layer 5	1486	[1441, 1521]
Layer 4	6	[2, 10]
Layer 3	26	[16, 33]
Layer 2	96	[80, 113]
Layer 1	342	[321, 366]
Layer 0	1067	[1067, 1067]

The table shows the results of experiment 1: creating random networks with fixed density, setting layer 0 firms at random, and computing the number of firms that belong to each production layer. The number of Monte-Carlo simulations is $M = 100$. The table shows the median number of firms in each layer across the simulations as well as the 95% confidence interval. Panels A, B and C show the results for $p = 3.61 \cdot 10^{-4}$, $1.44 \cdot 10^{-3}$, and $1.81 \cdot 10^{-4}$, respectively.

counterpart: (a) The empirical distribution of firms across layers is characterized by an almost monotonic distribution (i.e., a triangular shape). By contrast, the experiment-implied distribution is not triangular. The number of firms decreases from layer 0 to layer 4, and then spikes sharply in layer 5 (i.e., a skewed U-shaped distribution). This can be rationalized by the fact that we bundle all firms with vertical position 5 or higher into the top layer; (b) The number of firms in the top layer in the data is too small to be random. The lower bound of the number of firms in the top layer is 884 in the simulations versus 24 in the data.

In the data, the number of firms that fall into layer 5 is on average 24 (see Table 14). Panel B of Table OA.8 shows that in order for a random network to have the same number of firms in the top layer as in the data, the density p of the network has to be quadrupled. This density

is counterfactually large. While a larger density parameter allows to match the number of firms in layer 5, the distribution of firms in Panel B still diverts from the data. The number of firms increases from layer 0 to layer 1 and then drops from layer 2 to 4, violating monotonicity. Panel C shows that when p drops in relation to the benchmark, the experiment-implied distribution is still at odds with the data. In the simulations, the number of firms in layer 5 is larger than all other layers in sharp contrast to the data. This strictly violates a mechanical “triangular” property.

Overall, the vertical position computation implies an uneven distribution of firms across the production layers even for a random network. However, in all cases, it does not yield a triangular shape. We therefore cannot reject the hypothesis that the distribution in the data follows the economic notion of Hayekian triangle.

Despite the discrepancies between the distribution implied by a random network and the empirical distribution, the Monte-Carlo experiment suggests a robust feature. Fixing N_0 , in networks with the same density as the data or greater, the top layer has fewer firms compared to the bottom. This feature is common to the data. Whether this feature arises mechanically or not is only important for our empirical results if it can drive the TMB spread. Thus, it is important to rule out a mechanical connection between this feature and stylized fact 1 (the TMB spread). We do so using a separate Monte-Carlo experiment, described below.

Let M be the number of iterations. In each iteration m we do the following (experiment 2):

(1) For $t = 1$ to T :

(1.1) Create a vector u_t^m as a random shuffling of the vector $[1..\overline{N}]$.

(1.2) Set $VP_t^m(j) = VP_t(u_t^m)$, $\forall j \in \{1..\overline{N}\}$

(2) Sort firms every month t into portfolios based on VP_{t-1}^m

(3) Compute the implied $TMB(m)$ spread.

According to the algorithm above, we keep the number of firms in each layer identical to the empirical distribution, but we assign the vertical position to firms at random (reshuffled). In particular, the number of firms in the top and bottom layers of production is equal, month-by-month, to the data. If the (uneven) distribution of firms across the layers drives the TMB spread,

the confidence interval of the TMB spread across all M iterations should be positive. We set M to 1000 to ensure a sufficiently comprehensive confidence interval.

Table OA.9: **Monte-Carlo: Random Vertical Position Assignment and TMB**

	Equal-Weighting		Value-Weighting	
Layer 5	1.29	[0.72, 1.79]	0.91	[0.23, 1.53]
Layer 4	1.27	[0.96, 1.58]	0.85	[0.46, 1.25]
Layer 3	1.26	[1.10, 1.42]	0.80	[0.53, 1.02]
Layer 2	1.26	[1.19, 1.33]	0.78	[0.66, 0.89]
Layer 1	1.26	[1.18, 1.35]	0.78	[0.64, 0.91]
Layer 0	1.26	[1.19, 1.32]	0.78	[0.66, 0.89]
TMB (5-0)	0.02	[-0.54, 0.54]	0.12	[-0.54, 0.77]

The table shows the results of experiment 2: reshuffling the vertical position between firms in each month, over 1000 simulations. The table shows the median average return for each layer across the simulations, as well as the 95% confidence intervals for average returns. The sample period is from 2003/11–2013/02.

Table OA.9 reports the distribution of the average returns for each production layer across the M iterations. The median value-weighted TMB spread across M simulations is 12bps and the equal-weighted spread is 2bps per-month. In comparison to the empirical spread (over 100bps per month) these magnitudes are minuscule. Moreover, the lower bound of the experiment-implied confidence interval for TMB is negative. These results suggest that the distribution of firms across the layers, and the small number of top-layer firms, are immaterial for the TMB spread. Two related features in Table OA.9 also support this claim. First, the patterns of returns across the layers is, for the most part, flat. Second, across all simulations, the probability of obtaining a TMB spread as large as the one observed in the data is 0.9%. In sum, even if the vertical position computation in Equation (1) could mechanically result in a smaller number of firms in the top layer, this does not induce a positive TMB spread.

OA.10 Two Additional Alternative Explanations

In this section, we consider two additional alternative explanations for the TMB spread: (1) the lower persistence of vertical positions in upper layers; (2) a “bullwhip effect” due to larger forecasting errors in upper layers.

Persistence of Transitions Between Layers. Table OA.10 in Online Appendix reports monthly transition probabilities of firms across the layers during the sample period. The matrix shows that vertical positions are rather persistent at lower layers (98% for layer 1), but less persistent

Table OA.10: **Transition Matrix of the Vertical Position**

Layer in month t	Layer in month $t + 1$				
	Layer 5	Layer 4	Layer 3	Layer 2	Layer 1
Layer 5	0.83	0.07	0.06	0.04	0.01
Layer 4	0.02	0.88	0.06	0.04	0.00
Layer 3	0.00	0.02	0.92	0.05	0.01
Layer 2	0.00	0.00	0.01	0.97	0.01
Layer 1	0.00	0.00	0.00	0.01	0.98

This table presents the transition probability of a firm’s vertical position from one month to another. The matrix of transition probabilities is computed using monthly data from September 2003 to December 2012. Layer zero is not a part of the transition matrix because it does not change given that it is based on the time-invariant Global Industry Classification Standard (GICS) code reported in Compustat.

at the top layer (83%).⁸ The model assumes that a firm’s vertical position is fixed over time for tractability. In Section OA.5.4 of Online Appendix we assess quantitatively whether this difference in persistence can drive the spread. Using comparative statics that incorporate the transition probability matrix in Table OA.10, we show that allowing for dynamic transitions between layers reduces the TMB spread, but the effect is very small. The spread drops because the top layer can either remain in layer 5, or transition downstream. In the latter case its expected return must fall by the monotonic relation between vertical position and premia. However, because the model-implied TMB spread is slightly larger than the data, this downward adjustment of the spread makes the model-implied TMB identical to the empirical counterpart of 11.3%.

Bullwhip Effect. Due to forecasting errors that accumulate upward following a change in consumer demand, the top of the supply chain may face greater uncertainty in response to demand shocks than the bottom layer. This amplification in demand uncertainty is coined the “bullwhip effect” (Lee, Padmanabhan, and Whang (1997)). The bullwhip effect literature does not make a prediction regarding the risk premia of firms with different upstreamness. A potential implication of the bullwhip effect is that upstream firms should experience greater inventory turnover volatility. Consistently, in untabulated results we verify that inventory turnover volatility increases with the vertical position.⁹ However, greater inventory turnover volatility does not imply a higher risk premium in the data. When we sort firms into portfolios each quarter based on their lagged inventory turnover volatility, using a long sample from 1978 to 2013, the return spread between

⁸We note that if some supplier-customer relationships are not consistently covered in the data, it may cause a downward bias in our estimate of the vertical position persistence.

⁹We first compute inventory turnover for each firm in our sample, defined as the ratio of sales to the average stock of inventory in the current and the previous quarters. We then use the past 12 quarters to compute firms’ inventory turnover volatility.

the high and low inventory turnover volatility portfolios is either insignificant (value-weighted) or negative (equal-weighted).

OA.11 Supplemental Tables

This section contains supplemental tables mentioned in the paper.

Table OA.11: **Correlations of Returns across Layers**

	Layer 5	Layer 4	Layer 3	Layer 2	Layer 1	Layer 0
Layer 5	1.00					
Layer 4	0.67	1.00				
Layer 3	0.70	0.89	1.00			
Layer 2	0.68	0.87	0.90	1.00		
Layer 1	0.63	0.80	0.82	0.93	1.00	
Layer 0	0.57	0.70	0.70	0.84	0.89	1.00

This table reports the correlations between the returns of the value-weighted portfolios formed based on the vertical position measure defined in Equation (1).

Table OA.12: Risk Exposures to Aggregate Productivity: Model versus Data

	TMB	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
Panel A. $R_{i,t}^e = const + \beta_1 \Delta LaborProd_t + \beta_2 \Delta LaborProd_t^2 + error$							
$\beta_1 :$							
Data	8.044	2.484	4.117	6.609	9.654	10.513	10.528
Model	1.874	1.071	1.533	1.899	2.188	2.415	2.945
$\beta_2 :$							
Data	-1.354	-0.269	-0.596	-1.053	-1.511	-1.791	-1.623
Model	-1.497	0.157	-0.018	-0.645	-1.319	-1.928	-1.340
$\hat{\beta}_{LaborProd} = E[\frac{\partial R^e}{\partial \Delta Prod}] = \beta_1 + 2\beta_2 E[\Delta Prod]:$							
Data	3.254	1.530	2.006	2.882	4.308	4.175	4.784
Model	1.845	1.074	1.533	1.886	2.162	2.377	2.919
Panel B. $R_{i,t}^e = const + \beta_1 \Delta Solow_t + \beta_2 \Delta Solow_t^2 + error$							
$\beta_1 :$							
Data	2.863	-0.392	-0.352	0.142	0.681	0.217	2.472
Model	2.182	1.272	1.833	2.279	2.629	2.898	3.454
$\beta_2 :$							
Data	-0.362	0.387	0.365	0.299	0.192	0.283	0.025
Model	-1.134	0.480	0.471	0.046	-0.460	-0.958	-0.653
$\hat{\beta}_{Solow} = E[\frac{\partial R^e}{\partial \Delta Solow}] = \beta_1 + 2\beta_2 E[\Delta Solow]:$							
Data	1.937	0.599	0.583	0.909	1.171	0.942	2.536
Model	2.168	1.278	1.839	2.279	2.623	2.886	3.446

The table shows the results of the following regression: $R_{i,t}^e = const + \beta_1 \Delta X_t + \beta_2 \Delta X_t^2 + error$, where $R_{i,t}^e$ are layer excess portfolio returns, and X_t is an aggregate productivity measure. Panel A sets X to labor productivity, while panel B sets X to Solow residual. The model population moments are based on a simulated path of length 100,000 for layers' excess returns and for aggregate productivity. The overall exposures $\beta_{i,X}$ of layers portfolio returns to aggregate productivity X combine the linear term β_1 and the quadratic term β_2 . aggregate productivity growth. The data point estimates are taken from Table 2.

Table OA.13: **Impulse Responses: From Productivity to 1-Year Ahead Firms' Outcomes**

	Investment Rate		Sales/Assets	
	Model	Data	Model	Data
Top	0.21	0.24	0.20	0.22
Mid	0.14	0.21	0.18	0.29
Bottom	0.09	0.11	0.06	0.09
Diff	0.12	0.13	0.14	0.13

The table shows point estimates of the effect of one-standard deviation Cholesky shocks from productivity growth to 1-year ahead y_j , where y is either investment rate or sales-to-assets ratio, and $j \in \{Top, Mid, Bottom\} = \{4, 2, 0\}$. The impulse-responses are computed from a VAR that includes productivity growth, and the y variables of layers 4 to 0 in that order. Diff refers to the difference between Top and Bottom of each column. Productivity growth is measured as Solow residual growth in the data and as ΔZ_t in the model. For each firm, investment rate by the ratio of capital expenditure (CAPX) to capital (PPENT). In the data, we compute each layer's $y_{j,t}$ variable as the median across all firms in the same layer, measured at quarterly frequency over the benchmark sample period. In the model, $k_{j,t}$ and $i_{j,t}$ are based on a simulated path of 100,000 periods.

Table OA.14: **Sample Splits by Book-to-Market, Depreciation, Organization Capital, and Inventory**

	Value-weighted		Equal-weighted		Value-weighted		Equal-weighted	
	Panel A. Book-to-market split				Panel B. Depreciation split			
	Low	High	Low	High	Low	High	Low	High
Layer 5	13.50	15.91	5.21	19.56	18.95	10.81	13.55	12.14
Layer 4	10.05	11.39	1.49	11.37	12.65	8.90	8.65	3.96
Layer 3	4.40	9.06	0.64	9.75	7.51	4.12	8.51	1.91
Layer 2	3.97	8.86	0.46	7.93	6.63	2.57	6.47	2.02
Layer 1	4.88	5.45	3.14	4.77	5.80	3.13	4.13	3.68
Layer 0	5.32	4.03	2.14	3.51	4.94	4.79	2.87	3.06
TMB	8.18	11.87	3.07	16.05***	14.01*	6.02	10.68**	9.08**
t-stat	(1.23)	(1.59)	(0.68)	(3.08)	(1.92)	(1.27)	(2.2)	(2.07)
	Panel C. Organization capital split				Panel D. Inventory split			
	Low	High	Low	High	Low	High	Low	High
Layer 5	13.45	5.49	16.11	7.63	14.54	15.48	6.99	17.73
Layer 4	13.61	8.00	0.73	13.32	8.35	14.42	1.93	10.83
Layer 3	6.91	7.69	5.33	5.07	8.84	5.26	5.26	4.83
Layer 2	6.61	4.35	4.34	4.22	4.82	7.33	2.41	6.05
Layer 1	4.23	6.77	3.31	4.69	5.19	4.27	4.18	3.66
Layer 0	4.16	6.03	1.12	5.18	5.78	3.47	1.52	4.41
TMB	9.30**	-0.54	14.99***	2.44	8.76	12.01**	5.47	13.33**
t-stat	(2.34)	(-0.04)	(4.15)	(0.27)	(1.28)	(1.99)	(1.17)	(2.52)

This table reports the annualized continuously compounded excess returns by vertical position for different subsamples. Firms in each layer are split into two subsamples of equal size based on the book-to-market equity ratio (Panel A), depreciation rate (Panel B), ratio of organization capital to total assets (Panel C), or ratio of inventory to firm value (Panel D). All ratios are computed using the annual Compustat database. The results are based on a monthly sample from 2003/11–2013/02, aggregated over a rolling window of 12 months to form annualized returns. Newey-West t-statistics for the TMB portfolio are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table OA.15: Time series regressions of the TMB portfolio return: Equal-weighted

	(1)	(2)	(3)	(4)	(5)	(6)
Rm-Rf	-0.107 (-1.10)	-0.057 (-0.50)	0.027 (0.23)	0.031 (0.27)	-0.364*** (-3.00)	
SMB		-0.020 (-0.09)	-0.071 (-0.34)	0.135 (0.65)		
HML		-0.242 (-1.28)	-0.127 (-0.67)			
MOM			0.234** (2.53)			
I/A				-0.784*** (-2.69)		
ROE				0.504** (2.21)		
IMC					0.563*** (3.29)	
DMN						-0.112 (-1.57)
Constant	1.129** (2.62)	1.157*** (2.68)	1.114*** (2.64)	1.023** (2.45)	1.242*** (3.00)	1.050** (2.46)
R^2	0.011	0.026	0.081	0.137	0.100	0.022

This table presents the results of the time series regressions of the equal-weighted Top-Minus-Bottom (TMB) portfolio returns on various risk factors and the durable-nondurable spread. The TMB portfolio is constructed by taking a long position in the value-weighted portfolio of companies with a vertical position of five or above and shorting a value-weighted portfolio of companies with a vertical position of zero. The returns are computed from November 2003 to February 2013. The factor description and construction are identical to those in Table OA.6. T-statistics are reported in parentheses. Significance at the 1%, 5%, and 10% levels are indicated by ***, ** and * respectively.

Table OA.16: **Empirical Robustness of the Top-Minus-Bottom Spread**

	(1)	(2)	(3)	(4)	(5)
	Quarterly Rebalance	Annually Rebalance	Four-month Lag	Median Distance	Five Layers
Panel A. Return by vertical position					
Layer 5	1.90	1.85	1.62	1.50	-
Layer 4	1.31	1.05	1.24	1.00	1.43
Layer 3	1.00	0.95	0.98	1.04	0.99
Layer 2	0.87	0.88	0.81	0.84	0.87
Layer 1	0.73	0.72	0.68	0.58	0.73
Layer 0	0.73	0.73	0.71	0.73	0.73
Panel B. Spreads					
Top-minus-Bottom	1.17** (2.48)	1.12*** (2.94)	0.91* (1.97)	0.77* (1.77)	0.70* (1.73)
Top-minus-Layer 1	1.17*** (2.63)	1.13*** (3.39)	0.94** (2.18)	0.92** (2.26)	0.70** (2.04)

The table presents excess returns of the different layers (Panel A) and spreads (Panel B) under multiple alternative portfolio formation methods. The portfolios are constructed in an identical manner to the benchmark case, except the following changes: (1) sorting firms into portfolios based on their vertical position only at the end of each quarter (March, June, September and December); (2) sorting firms into portfolios based on their vertical position only once a year (at the end of June); (3) sorting firms into portfolio every month t based on the vertical position computed in month $t - 4$; (4) using the operator median instead of minimum in equation (1) for the vertical position assignment; and (5) grouping firms into only five layers, where the top layer includes all firms with vertical position of 4 or above. In columns (1)-(4), the top layer is layer five, and in column (5) it is layer 4. All numbers are in percentage units. T-statistics are reported in parentheses. Significance at the 1, 5, and 10 percent levels are indicated by ***, **, and *, respectively. The returns are computed from November 2003 to February 2013.

Table OA.17: Time series regressions of the TMB portfolio return: Annual Rebalancing

	(1)	(2)	(3)	(4)	(5)	(6)
Rm-Rf	0.017 (0.19)	0.105 (1.04)	0.229** (2.36)	0.187* (1.79)	-0.153 (-1.38)	
SMB		-0.367* (-1.94)	-0.443** (-2.53)	-0.245 (-1.33)		
HML		-0.012 (-0.07)	0.157 (0.99)			
MOM			0.346*** (4.49)			
I/A				-0.674** (-2.60)		
ROE				0.318 (1.57)		
IMC					0.371** (2.38)	
DMN						-0.167*** (-2.71)
Constant	1.106*** (2.89)	1.129*** (2.97)	1.066*** (3.03)	1.076*** (2.90)	1.181*** (3.13)	1.076*** (2.92)
R^2	0.000	0.034	0.187	0.129	0.050	0.062

This table presents the results of the time series regressions of the value-weighted Top-Minus-Bottom (TMB) portfolio returns on various risk factors and the durable-nondurable spread. As opposed to the benchmark case, firms are sorted into vertical position portfolios only once a year, and the end of June. The TMB portfolio is constructed by taking a long position in the value-weighted portfolio of companies with a vertical position of five or above and shorting a value-weighted portfolio of companies with a vertical position of zero. The returns are computed from November 2003 to February 2013. The factor description and construction are identical to those in Table OA.6. T-statistics are reported in parentheses. Significance at the 1%, 5%, and 10% levels are indicated by ***, ** and * respectively.

Table OA.18: **Top-Minus-Bottom Spread: Constant Portfolio Assignment**

Panel A. Monthly return by vertical position	
Layer 5	1.141
Layer 4	0.969
Layer 3	0.965
Layer 2	0.855
Layer 1	0.737
Layer 0	0.731
Panel B. Spread	
Top-minus-Bottom	0.409
	(0.920)

The table presents monthly returns for 11/2003–02/2013 of different layers (Panel A) and the TMB spread (Panel B) when firms are assigned into layer portfolios by their first vertical positions observed during the sample period. After the initial assignment, a firm remains in the same layer portfolio throughout the sample period. All numbers are in percentage units. T-statistics are reported in parentheses.