Aggregation in Heterogeneous-Firm Models: Theory and Measurement*

Yann Koby
Princeton University

Christian K. Wolf
Princeton University

July 15, 2020

Abstract: In structural models of investment with real and financial frictions to capital accumulation, the distribution of capital across firms is well-known to affect the sensitivity of partial equilibrium investment demand to economic stimulus. We show that general equilibrium price responses will undo this state dependence if investment is sufficiently price-elastic. While large price elasticities are a feature of most standard models of investment, we argue that they are rejected empirically: Quasi-experimental evidence on firm-level investment responses to tax changes suggests price elasticities that are orders of magnitude too small to generate meaningful general equilibrium smoothing. In models with data-consistent small price elasticities, the aggregate effects of monetary and fiscal stimulus are indeed sensitive to the cross-sectional distribution of firm capital and so to broader business-cycle conditions. Finally, we show that our results apply without change to the dynamics of aggregate durable consumption.

Keywords: Firm heterogeneity, lumpy investment, financial frictions, aggregation, state dependence, durable consumption. *JEL codes:* D21, D22, E22, E62.

*Email: `ykoby@princeton.edu` and `ckwolf@princeton.edu`. We received helpful comments from Mark Aguiar, Ben Moll, Greg Kaplan, Mikkel Plagborg-Møller, Gianluca Violante, Tom Winberry, and seminar participants at several venues. The first draft of this paper was written while Wolf was visiting the European Central Bank, whose hospitality is gratefully acknowledged.

1 Introduction

A voluminous theoretical literature debates whether the empirically well-documented lumpiness of firm-level investment also matters for aggregate investment dynamics.[1] In partial equilibrium, the answer is yes: the more firms plan to make an extensive margin investment, the more responsive is total investment to any incremental stimulus. In general equilibrium, this state dependence may or may not disappear: Previous work shows that two structural models, both featuring lumpy firm-level investment, can dramatically disagree on the extent to which general equilibrium price responses smooth out any such partial equilibrium state dependence (Khan & Thomas, 2008; Winberry, 2018). If general equilibrium smoothing is strong enough, then aggregate investment behaves as in a simpler economy with a standard neoclassical firm block – the production side aggregates, and the cross-sectional distribution of capital holdings becomes irrelevant for aggregate investment dynamics. General equilibrium price effects are similarly known to mute the importance of firm-level financial frictions (Zetlin-Jones & Shourideh, 2017; Khan & Thomas, 2013).

We make three contributions to the aggregation debate. First, we show that general equilibrium price effects undo partial equilibrium state dependence only if firm investment is sufficiently sensitive to changes in the cost of capital. With price-sensitive investment, shifts in the cross-sectional distribution of capital holdings affect investment demand *given* prices, but are easily smoothed out in general equilibrium through small *changes* in prices. The literature so far disagreed on aggregate state dependence precisely because of implicit disagreement on this price elasticity. Second, we directly measure investment price elasticities using quasi-experimental evidence on the firm-level response of investment to transitory tax changes. Our preferred estimate of the average interest rate semi-elasticity of investment is around five per cent, multiple orders of magnitude below that in popular models with strong general equilibrium smoothing. Third, we document that, in calibrated heterogeneous-firm models consistent with (i) standard real and financial investment frictions and (ii) our measured (low) price elasticities, the aggregate effects of monetary and fiscal stimulus depend sensitively on the cross-sectional distribution of capital and so broader business-cycle conditions. Finally, we note that our theoretical results on general equilibrium smoothing apply without change to the dynamics of aggregate durables consumption.

To make our first point, we study state dependence in the response of aggregate invest-

[1] For example see Caballero & Engel (1999), Thomas (2002), Khan & Thomas (2008), Bachmann et al. (2013) and Winberry (2018).

ment to macro shocks in an infinite-horizon business-cycle model with aggregate risk, general enough to nest most popular structural models of investment (e.g. Khan & Thomas, 2008, 2013; Winberry, 2018) as well as recent contributions to the broader business-cycle literature (Smets & Wouters, 2007; Justiniano et al., 2010). In variants of this model with real and financial frictions to capital reallocation, the distribution of firms across their idiosyncratic state space – in particular their capital holdings – is well-known to shape the response of investment to shocks in partial equilibrium, ignoring any general equilibrium price feedback (Caballero & Engel, 1999; Khan & Thomas, 2013).

We show analytically that the extent to which any partial equilibrium state dependence is smoothed out in general equilibrium is, to first order, governed by local price elasticities of investment demand and supply. To build intuition for the role of investment price elasticities in shaping general equilibrium smoothing, we consider a simplified variant of our economy in which an exogenously fixed fraction $\xi \in (0, 1)$ of firms is not allowed to adjust their capital stock – a reduced-form stand-in for time variation in the partial equilibrium sensitivity of investment to shocks. We show that, if either firm revenue functions are linear in the scale of production, or if capital goods are infinitely-lived, then all aggregate prices and quantities are independent of ξ. In either case, investment is infinitely sensitive to changes in the cost of capital around the equilibrium price path of a standard neoclassical economy with $\xi = 0$. As a result, arbitrarily small changes in real interest rates are enough to neutralize any partial equilibrium shifts in firm investment demand due to exogenous changes in ξ. In a static model, this intuition is easily visualized: Infinite price elasticities correspond to flat investment demand curves, but horizontal shifts of a flat demand curve have no effect on prices or quantities, independently of the slope of investment supply. We show that, in our infinite-horizon economy, the analogue of this "flat demand curve" logic is rank deficiency in a matrix $\mathcal{H}$ of dynamic investment demand and supply elasticities, with ξ-induced shifts of aggregate investment demand (at fixed prices) lying in the null space of $\mathcal{H}$.

Our theoretical aggregation result rationalizes disagreement in previous work. In Khan & Thomas (2008), the (annual) partial equilibrium interest rate semi-elasticity of firm-level investment is almost 500 per cent, and so small changes in the cost of capital easily undo shifts in investment demand due to changes in the cross-sectional distribution of firm capital holdings. For example, even if investment demand were to respond by 10 percentage points more in a brisk boom than a deep recession, interest rates would need to only drop by 2 *basis points* to fully offset the partial equilibrium asymmetry. In the model of Winberry (2018), that same interest rate elasticity is around 7 per cent, so similar general equilibrium smoothing would require interest rates to move by almost 2 percentage points – a change that itself would elicit a strong response of investment supply.

We next provide empirical discipline for investment price elasticities. Disagreement in previous work could arise because standard model estimation targets – often distributions of investment rates and their persistence – are simply not very informative about the interest rate elasticity of investment, and so tell us little about the strength of general equilibrium smoothing. We instead provide direct empirical discipline through detailed firm-level evidence on the response of investment to tax stimulus. Similar to Zwick & Mahon (2017), we estimate firm-level investment responses to the bonus depreciation episodes of 2001-2004 and 2008-2010. Importantly, we establish a formal connection between the econometric estimand of such regressions and the interest rate elasticity of investment that we have identified as central to the strength of general equilibrium smoothing.[2] Our estimates correspond to an average semi-elasticity of investment with respect to changes in the cost of capital of around five per cent, around two orders of magnitude smaller than the elasticity in Khan & Thomas (2008). Firm-level evidence thus strongly favors structural models with downward-sloping aggregate investment demand curves, rejecting the strong general equilibrium smoothing required to undo partial equilibrium state dependence.

We illustrate the practical implications of our results in two structural models, each featuring (i) partial equilibrium state dependence due to real and financial investment frictions and (ii) weak general equilibrium smoothing due to small investment price elasticities. First, we study monetary policy in a model with lumpy investment. As usual, fixed investment adjustment costs imply that the partial equilibrium elasticity of investment with respect to interest rate changes is state-dependent; in particular, it is lower in a TFP-induced recession. Given our muted investment price elasticities, this asymmetry survives in general equilibrium, and so monetary policy "pushes on a string" in recessions. Second, we study fiscal policy in a model with firm-level financial frictions in the form of an earnings-based borrowing constraint. In this model, expansionary fiscal policy is particularly effective in financial recessions: A fiscal spending expansion boosts firm earnings, pushing up investment of financially constrained firms, and so crowding-*in* aggregate investment.

All of our results apply without change to state dependence in the response of household durables consumption to economic stimulus. If either household preferences over durables

[2] While infinite *interest rate* elasticities are central to our aggregation result, general equilibrium feedback is generally shaped by the response of investment to changes in all kinds of aggregate prices, including wages and sales prices. We show, using standard investment theory, that one elasticity – the interest rate elasticity – is robustly informative about all of them.

are linear or if durables are infinitely-lived, then the interest rate elasticity of durables spending is infinite, and so the distribution of durables holdings across households is irrelevant for aggregate consumption dynamics. In Berger & Vavra (2015) and McKay & Wieland (2019), depreciation rates are high and household preferences are far from linear, so general equilibrium smoothing is weak, and aggregate stimulus is state-dependent.

Literature. Our work relates to three main strands of literature.

First, we contribute to the long literature on how investment decisions at the micro level shape aggregate investment dynamics. In this literature, two particular frictions – non-convex firm-level adjustment costs and financial (leverage) constraints – have received most attention. Caballero & Engel (1999) argue that fixed costs of investment can lead to a time-varying sensitivity of investment to shocks; in recessions, firms are roughly at their target capital, so they are particularly reluctant to respond to any incremental stimulus. Khan & Thomas (2008) show that, while such non-convexities invariably play an important role in partial equilibrium and at the firm level, they may actually not survive to shape aggregate investment in general equilibrium. In contrast, Bachmann et al. (2013) and Winberry (2018) present models in which micro lumpiness does matter for aggregate investment dynamics. The extent to which firm-level financial frictions affect aggregate investment is similarly contested in previous work (Zetlin-Jones & Shourideh, 2017; Khan & Thomas, 2013).

Second, we relate to previous work on the aggregation properties of heterogeneous-agent models. The closest precursor to our analysis is House (2014) who – in a deterministic partial equilibrium model with fixed adjustment costs – shows that investment demand curves for long-lived capital goods are flat.[3] We offer additional insights by emphasizing the generality of the aggregation logic: Our results apply (i) in a large space of general equilibrium models with aggregate risk, (ii) for generic shifters of partial equilibrium investment demand, and also (iii) to household consumption of durables. Relative to the discussion in Winberry (2018), we clarify that it is the price elasticity of investment – rather than the cyclicality of real interest rates per se – that governs the strength of general equilibrium smoothing. Other well-known aggregation results for heterogeneous-agent models rely on the linearity of policy functions in idiosyncratic state variables (e.g. Gorman, 1961; Werning, 2016). In our case, policy functions are clearly not linear in firm states – and so investment *given prices*

[3] His analysis of aggregate equilibria is restricted to the investment market. With a full general equilibrium model closure, however, the shifter Z_t in his reduced-form investment supply function itself depends on firm behavior; as a result, the investment market alone does not afford a full characterization of general equilibrium dynamics. Our perfect foresight equilibrium decomposition offers precisely such a characterization.

is affected by the cross-sectional state distribution –, but large price elasticities allow this state dependence to be smoothed out in general equilibrium through *changes* in prices.

Third, we leverage and extend recent empirical evidence on the price elasticity of firm investment. A long literature has tried to exploit cross-sectional variation to study the effect of tax policy on investment (Cummins et al., 1994; Goolsbee, 1998; House & Shapiro, 2008; Zwick & Mahon, 2017). Among those papers, our empirical analysis most closely builds on Zwick & Mahon, who study the response of investment to the bonus depreciation episodes of 2001-2004 and 2008-2010. It is widely regarded as a defect of such cross-sectional analyses that, by construction, general equilibrium effects are differenced out (e.g. Wolf, 2019). For our purposes, however, partial equilibrium elasticities – the slope of investment demand – is precisely what is needed to learn about the strength of general equilibrium smoothing. The overall approach of using micro data to discipline general equilibrium adjustment closely follows Auclert et al. (2018), who leverage data on household consumption behavior to learn about aggregate Keynesian multiplier effects.

Outlook. Section 2 presents our theoretical argument for the role of price elasticities in shaping the strength of general equilibrium smoothing. Section 3 shows that existing calibration strategies for heterogeneous-firm models are largely silent on these price elasticities, and so instead uses microeconomic quasi-experiments for more direct discipline. Section 4 then illustrates through two examples how, with data-consistent (low) price elasticities, the aggregate effects of investment stimulus become state-dependent. Section 5 concludes, and supplementary details as well as all proofs are relegated to several appendices.

2 Aggregation and the price elasticity of investment

We show that the interest rate elasticity of investment plays a central role in shaping the extent to which partial equilibrium state dependence in investment demand shapes the response of aggregate investment to economic stimulus. Section 2.1 outlines a family of models to which our arguments apply. We present a simple analytical aggregation result in Section 2.2, and then in Section 2.3 use it to rationalize disagreement in previous work.

2.1 Model

Time is discrete and runs forever. The model is populated by households, firms, and a government, and shocks to aggregate productivity are the single source of aggregate risk.[4]

We study first-order perturbations starting from an arbitrary initial aggregate state $s_0 \in \mathcal{S}$, where $\mathcal{S}$ denotes the aggregate state space, and s_0 is a typical element. This solution concept allows us to account for state dependence in an economy with aggregate risk in a computationally and analytically convenient way.[5] We use s_0 subscripts to indicate dependence on the initial state, but suppress this dependence whenever there is no risk of confusion. Following Fernández-Villaverde et al. (2016) and Boppart et al. (2018), we characterize impulse responses associated with our first-order perturbation solution as perfect foresight transition paths for vanishingly small, one-time unexpected innovations ("MIT shocks") at time 0. The realization of a variable x at time t along the transition path will be denoted x_t, while the entire time path will be denoted $\mathbf{x}$. Throughout, hats denote deviations from the underlying deterministic transition path induced by the initial state, either at a given point in time $\hat{x}_t$ or for the entire transition path $\hat{\mathbf{x}}$. Finally, bars indicate values at the economy's (aggregate) deterministic steady state.

Production. Production and investment are undertaken by a unit continuum of perfectly competitive intermediate goods producers $j \in [0,1]$. They hire labor at spot wage rate w_t, sell their output at price p_t^I, buy investment goods at price q_t, and produce with the production function $y_{jt} = z_t(k_{jt-1}^{\alpha}\ell_{jt}^{1-\alpha})^{\nu}$, where z_t denotes aggregate TFP. Dividends are paid out to households, and discounted at the equilibrium real interest rate r_t, equal to the stochastic discount factor of the owner-households. Given a path of prices $\mathbf{p} = (\mathbf{r}, \mathbf{p}^I, \mathbf{w}, \mathbf{q})$ and productivity $\mathbf{z}$, the time-0 problem of an intermediate goods producer j is as follows:

$$\max_{\{d_{jt}, y_{jt}, \ell_{jt}, k_{jt}, i_{jt}, b_{jt}^f\}} \quad \mathbb{E}_0\left[\sum_{t=0}^{\infty}\left(\prod_{q=0}^{t-1}\frac{1}{1+r_q}\right)\zeta^t d_{jt}\right] \tag{1}$$

[4] None of our arguments about aggregation will hinge on the source of aggregate risk. We consider TFP shocks because they feature prominently in the analyses of Khan & Thomas (2008) and Winberry (2018).

[5] Formally, we study perturbations around the full deterministic transition path induced by the (arbitrary) initial state. If the economy is initially in its deterministic steady state, then our solution coincides with ordinary perturbations around that steady state.

such that

$$
\begin{aligned}
d_{jt} &= \underbrace{p_t^I y_{jt} - w_t \ell_{jt}}_{\pi_{jt}} - q_t i_{jt} - \phi(k_{jt}, k_{jt-1}) + b_{jt} - (1 + r_t^b) b_{jt-1} \\
y_{jt} &= z_t e_{jt} (k_{jt-1}^{\alpha} \ell_{jt}^{1-\alpha})^{\nu} \\
i_{jt} &= k_{jt} - (1-\delta) k_{jt-1} \\
b_{jt} &\leq \Gamma(q_t k_{jt-1}, \pi_{jt}) \\
d_{jt} &\geq \underline{d}
\end{aligned}
$$

Firms die at rate $1-\zeta$, face time-varying idiosyncratic productivity risk e_{jt}, invest and hire labor, pay out dividends, buy capital at price q_t, and raise risk-free debt b_{jt}. However, their investment opportunities are subject to a rich set of real and financial frictions: Investment incurs the (potentially non-convex) real adjustment cost $\phi(k_{jt}, k_{jt-1})$, equity issuance is constrained by a lower bound on dividend payments, and maximal firm leverage is limited, subject to either a collateral constraint on $q_t k_{jt-1}$ or a cash flow constraint based on firm revenues π_{jt}. The production block is general enough to nest most previous contributions to the quantitative heterogeneous-firm investment literature, including Khan & Thomas (2008, 2013) and Winberry (2018).

We present a recursive characterization of the firm problem in Appendix A.2. An initial condition is a distribution μ_0 of firms across their idiosyncratic state space $\mathcal{S}^f$, with typical element $s^f = (e, k, b)$. The distribution μ_0 is part of the initial aggregate s_0 of the economy. Aggregating across all firms j, we obtain aggregate investment demand, labor demand, and output supply functions $\mathbf{i} = \mathbf{i}(\mathbf{p}; \mathbf{z})$, $\boldsymbol{\ell}^d = \boldsymbol{\ell}^d(\mathbf{p}; \mathbf{z})$ and $\mathbf{y} = \mathbf{y}(\mathbf{p}; \mathbf{z})$.

Rest of the economy. Rather than explicitly characterizing the rest of the economy in terms of economic fundamentals, we instead make the following high-level assumption:

Assumption 1. *Given the initial state s_0, there exists a function $\mathbf{p} = \mathbf{p}(\mathbf{y} - \mathbf{i})$, mapping sequences of net output supply $\tilde{\mathbf{y}} \equiv \mathbf{y} - \mathbf{i}$ into sequences of aggregate prices $\mathbf{p}$ such that all actors in the (non-production) rest of the economy behave optimally if and only if $\mathbf{p} = \mathbf{p}(\tilde{\mathbf{y}})$.*

This assumption, while somewhat peculiar at first sight, is surprisingly unrestrictive. In Appendix A.1, we prove that even rich medium-scale DSGE models are consistent with Assumption 1. In particular, we show that many of the recent contributions to the structural heterogeneous-firm investment literature consider model closures consistent with this assumption (Khan & Thomas, 2008, 2013; Winberry, 2018; Ottonello & Winberry, 2018).

Our results on the role of investment price elasticities in general equilibrium smoothing do not depend on the specifics of the mapping $\mathbf{p}(\bullet)$. Ultimately, aggregation turns out to be a property of the production side of the economy, and so in particular is invariant to the preferred general equilibrium model closure.

Equilibrium. We can now formally define perfect foresight transition equilibria.

Definition 1. *Given an initial state s_0 and an exogenous aggregate TFP path $\{z_t\}_{t=0}^{\infty}$, a perfect foresight transition path equilibrium is a sequence of aggregate quantities $\{\tilde{y}_t, y_t, i_t, \ell_t\}_{t=0}^{\infty}$ and prices $\{r_t, p_t^I, w_t, q_t\}_{t=0}^{\infty}$ such that:*

1. Firm Optimization. *Given prices and aggregate TFP, the paths of aggregate investment $\mathbf{i} = \mathbf{i}(\mathbf{p}; \mathbf{z})$, labor demand $\boldsymbol{\ell} = \boldsymbol{\ell}(\mathbf{p}; \mathbf{z})$ and production $\mathbf{y} = \mathbf{y}(\mathbf{p}; \mathbf{z})$ are consistent with optimal intermediate goods producer behavior.*

2. Rest of the Economy. *Aggregate prices satisfy*

$$\mathbf{p} = \mathbf{p}(\tilde{\mathbf{y}})$$

3. Aggregate Consistency. *Net output supply equals net output demand,*

$$y_t - i_t = \tilde{y}_t$$

for all $t = 0, 1, 2, \ldots$.

2.2 A simple aggregation result

To build intuition for the role of investment price elasticities in shaping the extent of general equilibrium smoothing in general, and aggregation results in particular, we first consider a simpler version of our benchmark economy. The non-production side of the economy continues to be summarized by Assumption 1. The problem of intermediate goods producers, however, is simplified to feature (i) no heterogeneity in idiosyncratic productivity ($e_{jt} = \bar{e}$ for all j and t), (ii) no real adjustment costs ($\phi(\bullet, \bullet) = 0$) and (iii) no financial constraints on debt or equity issuance. However, we impose that a fraction $\xi \in (0, 1)$ of firms is forced to keep their capital stock fixed at $\bar{k}$ forever; intuitively, a friction of this sort mimics the effects of fixed investment adjustment costs in conjunction with time-varying cross-sectional distributions of capital.

Irrelevance. Given any path of prices $\mathbf{p}$ and aggregate productivity $\mathbf{z}$, the initial distribution of firms μ_0 over their idiosyncratic state space – that is, the fraction of non-adjusting firms ξ – affects overall firm investment demand:

$$\mathbf{i}_{s_0(\xi)}(\mathbf{p};\mathbf{z}) - \mathbf{i}_{\bar{s}}(\mathbf{p};\mathbf{z}) = -\xi \times \mathbf{i}_{\bar{s}}(\mathbf{p};\mathbf{z}) \tag{2}$$

For example, the larger is ξ, the smaller is the *partial equilibrium* increase in investment demand following a transitory increase in productivity z. Nevertheless, under some further assumptions on production technologies, aggregate *general equilibrium* prices and quantities turn out to be independent of ξ.

Proposition 1. *Suppose that either $\nu = 1$ or $\bar{r} + \delta = 0$, and that the equilibrium in Definition 1 for the special case with exogenous ξ exists and is unique. Then the equilibrium price paths $\mathbf{p}$ as well as all equilibrium aggregates – total consumption $\mathbf{c}$, total investment $\mathbf{i}$, total labor hired $\boldsymbol{\ell}$ and total output $\mathbf{y}$ – are independent of ξ, so the initial distribution of firms μ_0 over their idiosyncratic state space is irrelevant.*

Mechanism. The economic logic underlying the aggregation result in Proposition 1 is simple: In either limit $\nu \to 1$ or $\bar{r} + \delta \to 0$, investment demand of unconstrained (adjusting) firms is infinitely responsive to changes in the cost of capital. To first order:

$$\frac{di_j}{\bar{i}} = -\frac{1}{1-\nu}\frac{1}{\bar{r}+\delta}\frac{1}{\delta}dr \tag{3}$$

In our infinite-horizon economy, the dynamic *path* of investment demand is infinitely elastic around the general equilibrium price path of an economy with $\xi = 0$. Thus, while the initial distribution of firms μ_0 matters for investment demand *given prices*, arbitrarily small *changes in prices* are enough to bring all aggregate quantities back in line with those in an economy without any constrained firms – the result in Proposition 1.

In a static market of investment supply and demand, this intuition would be straightforward to visualize: Horizontal shifts of a flat investment demand curve have no effects on aggregate prices or quantities. Assumption 1 allows us to straightforwardly generalize this intuition to our infinite-horizon economy. Here, a price path $\mathbf{p}$ is part of a perfect foresight equilibrium if and only if

$$\mathbf{y}(\mathbf{p}(\tilde{\mathbf{y}});\mathbf{z}) - \mathbf{i}(\mathbf{p}(\tilde{\mathbf{y}});\mathbf{z}) = \tilde{\mathbf{y}} \tag{4}$$

(4) is a fixed-point relation with rich economic content: An equilibrium is a path of aggregate

prices – interest rates, wages, and so on – at which net output supply and demand coincide at all points in time. Building on the equilibrium characterization in (4), Proposition 2 shows how to formalize the static "flat demand curve" logic in our infinite-horizon economy with aggregate risk.

Proposition 2. *To first order, the perfect foresight transition path of prices satisfies*

$$\hat{\mathbf{p}}_{s_0(\xi)} \quad = \quad \hat{\mathbf{p}}_{\bar{s}} \quad + \quad \underbrace{\mathcal{H}}_{\text{GE Adjustment}} \quad \times \quad \underbrace{\left[\hat{\tilde{\mathbf{y}}}_{s_0(\xi)}(\mathbf{p}_{\bar{s}};\mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}}(\mathbf{p}_{\bar{s}};\mathbf{z})\right]}_{\text{Excess Demand/Supply}} \tag{5}$$

where $\mathcal{H} = \frac{\partial \mathbf{p}}{\partial \tilde{\mathbf{y}}} \times (I - \mathcal{G})^{-1}$, I is the infinite-dimensional generalization of an identity matrix,

$$\mathcal{G} \quad \equiv \quad \underbrace{\begin{pmatrix} \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{r}} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}^I} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{w}} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{q}} \end{pmatrix}}_{\text{Supply Elasticity}} \quad \times \quad \underbrace{\begin{pmatrix} \frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{p}^I}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{w}}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{q}}{\partial \tilde{\mathbf{y}}} \end{pmatrix}}_{\text{Inverse Demand Elasticity}} \tag{6}$$

and $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{r}}$, $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}^I}$, $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{w}}$ and $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{q}}$ are infinite-dimensional linear maps of price elasticities, evaluated at s_0. For either $\nu = 1$ or $\bar{r} + \delta = 0$, the map $\mathcal{H}$ is column rank-deficient, with

$$\left\{\hat{\tilde{\mathbf{y}}}_{s_0(\xi)}(\mathbf{p}_{\bar{s}};\mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}}(\mathbf{p}_{\bar{s}};\mathbf{z}) \in \mathit{null}(\mathcal{H})\right\}$$

At equilibrium prices of the economy with $\xi = 0$, the presence of constrained firms affects aggregate investment demand and output supply, so generically $\hat{\tilde{\mathbf{y}}}_{s_0(\xi)}(\mathbf{p}_{\bar{s}};\mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}}(\mathbf{p}_{\bar{s}};\mathbf{z}) \neq \mathbf{0}$ – the infinite-horizon analogue of a horizontal investment demand shift in a static economy. The shift in partial equilibrium excess demand is then translated into general equilibrium prices through the matrix $\mathcal{H}$, which combines intertemporal elasticities of net output supply and demand. In either limit $\nu \to 1$ or $\bar{r} + \delta \to 0$, investment price elasticities diverge, $\mathcal{H}$ becomes rank-deficient, and ξ-induced shifts of partial equilibrium excess demand lie in the kernel of $\mathcal{H}$ – the irrelevant horizontal shift of a flat demand curve.

The general equilibrium adjustment implicit in (5) is depicted graphically in Figure 1, which presents the impulse response of investment to an aggregate productivity shock in two particular parametric versions of our simple economy. The left panel shows that, with strongly curved firm revenue functions ($\nu = 0.5$) and so small investment price elasticities, the presence of a large fringe $\xi = 0.5$ of non-adjusting firms materially dampens the response of the macro-economy to an expansionary productivity shock. At equilibrium prices of the alternative economy with $\xi = 0$, investment by construction only increases by half as much with $\xi = 0.5$, and so markets do not clear. To restore market-clearing, prices adjust –

notably, real interest rates fall, leading to more investment and crowding out consumption. At new equilibrium prices, total investment is higher than at $\hat{\mathbf{p}}_{\bar{s}}$, but not as high as in the frictionless economy. The right panel shows that, with constant returns to scale and so large price elasticities, a similar initial net excess supply gap at $\hat{\mathbf{p}}_{\bar{s}}$ is inconsequential in general equilibrium. Consistent with the intuition offered above, an arbitrarily small change in prices is enough to induce unconstrained firms to adjust their investment, thus fully offsetting the inaction of constrained firms.

(a) Strong Curvature: $\nu = 0.5$ **(b)** Aggregation: $\nu = 1$

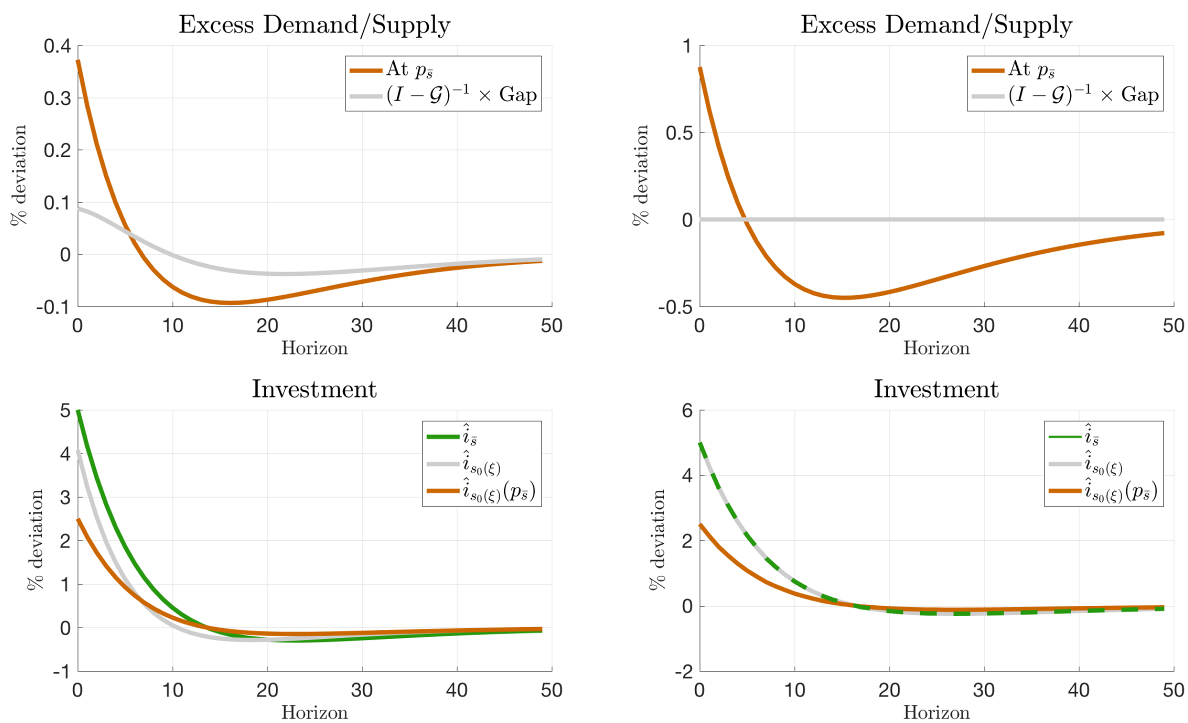

Figure 1: Aggregation along perfect foresight transition paths in infinite-horizon economies with lumpy firm-level investment. Model details are relegated to Appendix A.1.

Previous work. Our aggregation result extends House (2014). He studies a partial equilibrium model of investment demand and supply, and shows that – even in the presence of fixed adjustment costs – investment demand for long-lived capital goods is infinitely price-elastic. We show that elasticities also diverge as firm revenue functions become linear, and then prove an analytical general equilibrium aggregation result for a large family of infinite-horizon structural business-cycle models with aggregate risk. Our proof technique –

characterization of perfect foresight transition paths in sequence space – builds on similar approaches in Boppart et al. (2018), Auclert et al. (2019) and Wolf (2019). In particular, the general equilibrium adjustment matrix $\mathcal{H}$ at the heart of our equilibrium characterization is the firm-side equivalent of the intertemporal Keynesian cross matrix in Auclert et al. (2018).

The economic logic underlying the irrelevance of the cross-sectional firm distribution over idiosyncratic states is materially different from that in most standard heterogeneous-agent aggregation theorems (e.g. Gorman, 1961; Werning, 2016). As (2) makes clear, firm policy functions are *not* linear in idiosyncratic states, so aggregation does not obtain because average firm behavior is unaffected by the cross-sectional distribution for any given set of prices. Thus, in partial equilibrium, the firm block of our economy never aggregates. Instead, large elasticities to changes in prices mean that state dependence is easily smoothed out in general equilibrium through negligible aggregate price feedback. Aggregation results of this sort have important antecedents in earlier work on finance market microstructure, where the presence of deep-pocket, risk-neutral, rational investors pins down prices, even in the presence of irrational noise traders (Brunnermeier, 2001).

Extensions. Our simple aggregation logic applies beyond the simple model of Proposition 1. First, we show in Appendix A.1 that an analogous aggregation result obtains even when Assumption 1 is relaxed. Ultimately, aggregation is a property of our production block embedded into *some* general equilibrium closure, but the details of this closure do not matter.[6] Second, our choice of friction imposed on constrained firms – the inability to adjust capital holdings – was motivated by previous theoretical work (notably Caballero & Engel, 1999), but is no way essential for the aggregation result. For example, the production block would also aggregate if a random subset of firms was constrained to at most invest its own cash flow, as in Lian & Ma (2018). And third, Appendix A.3 proves an analogous aggregation result for the dynamics of aggregate consumption in a model with durables and non-durables, and where a random subset of households cannot adjust their durables holdings.

2.3 Aggregation in quantitative models

The logic of our analytical aggregation result in Proposition 2 carries through with little change to the bigger quantitative model of Section 2.1. We first discuss a characterization

[6] Formally, the details do not matter except for knife-edge special cases that ensure that all prices remain fixed also in general equilibrium. An example of such a special case are risk-neutral households.

of equilibrium transition paths exactly analogous to the simpler model, and then show how our results can rationalize disagreement in previous work.

Equilibrium characterization. As in the simpler model, a price path $\mathbf{p}$ is part of a perfect foresight equilibrium if and only if

$$\mathbf{y}(\mathbf{p}(\tilde{\mathbf{y}});\mathbf{z}) - \mathbf{i}(\mathbf{p}(\tilde{\mathbf{y}});\mathbf{z}) = \tilde{\mathbf{y}} \tag{7}$$

From (7), we straightforwardly arrive at the following characterization of state dependence in the response of aggregate prices to macroeconomic shocks.

Proposition 3. *Consider the structural model of Section 2.1. To first order, the perfect foresight transition path of prices satisfies*

$$\hat{\mathbf{p}}_{s_0} \quad = \quad \hat{\mathbf{p}}_{\bar{s}} \quad + \quad \underbrace{\mathcal{H}}_{\text{GE Adjustment}} \quad \times \quad \underbrace{\left[\hat{\tilde{\mathbf{y}}}_{s_0}(\mathbf{p}_{\bar{s}};\mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}}(\mathbf{p}_{\bar{s}};\mathbf{z})\right]}_{\text{Excess Demand/Supply}} \tag{8}$$

where $\mathcal{H} = \frac{\partial \mathbf{p}}{\partial \mathbf{c}} \times (I - \mathcal{G})^{-1}$,

$$\mathcal{G} \quad \equiv \quad \underbrace{\begin{pmatrix} \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{r}} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}^I} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{w}} & \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{q}} \end{pmatrix}}_{\text{Supply Elasticity}} \quad \times \quad \underbrace{\begin{pmatrix} \frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{p}^I}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{w}}{\partial \tilde{\mathbf{y}}} & \frac{\partial \mathbf{q}}{\partial \tilde{\mathbf{y}}} \end{pmatrix}}_{\text{Inverse Demand Elasticity}} \tag{9}$$

and $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{r}}$, $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}^I}$, $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{w}}$ *and* $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{q}}$ *are infinite-dimensional linear maps of price elasticities, evaluated at* s_0.

In rich models of firm-level real and financial frictions, the distribution of firms across idiosyncratic states is well-known to affect investment demand given prices (Caballero & Engel, 1999; Khan & Thomas, 2013), and so in particular changes firm net output supply $\hat{\tilde{\mathbf{y}}}$. Thus, analogously to our simple economy with the reduced-form friction $\xi > 0$, the net excess demand/supply term in (8) is generically non-zero. The extent to which any such state dependence given prices also survives to shape general equilibrium outcomes is again governed by the matrix $\mathcal{H}$, and so by the slope of firm investment demand functions. Thus, irrespective of the details of the underlying firm block, price elasticities are a firm-side "sufficient statistic" for the strength of general equilibrium smoothing. In particular, if firm net output supply at the initial state s_0 is sufficiently elastic around the equilibrium price path implied by the $\bar{s}$-economy, then aggregate price and quantity paths are identical in the two economies, irrespective of any state dependence *given* prices.

Relation to previous work. Our theoretical results can rationalize disagreement in previous work. Khan & Thomas (2008) study a standard business-cycle model augmented to feature fixed costs of investment, and document that the cross-sectional distribution of firm capital holdings only matters in partial equilibrium, but becomes quantitatively irrelevant in general equilibrium. More recently, Bachmann et al. (2013) and Winberry (2018) have presented structural models with much weaker general equilibrium smoothing. To understand the origins of this disagreement, we have replicated the models in Khan & Thomas (2008) and Winberry (2018), and in Figure 2 plot the (partial equilibrium) elasticities of investment today and in the future to an unexpected change in real interest rates today – that is, the first column of $\frac{\partial \mathbf{I}}{\partial \mathbf{r}}$.[7]

Investment Elasticities: Khan & Thomas (2008) vs. Winberry (2018)

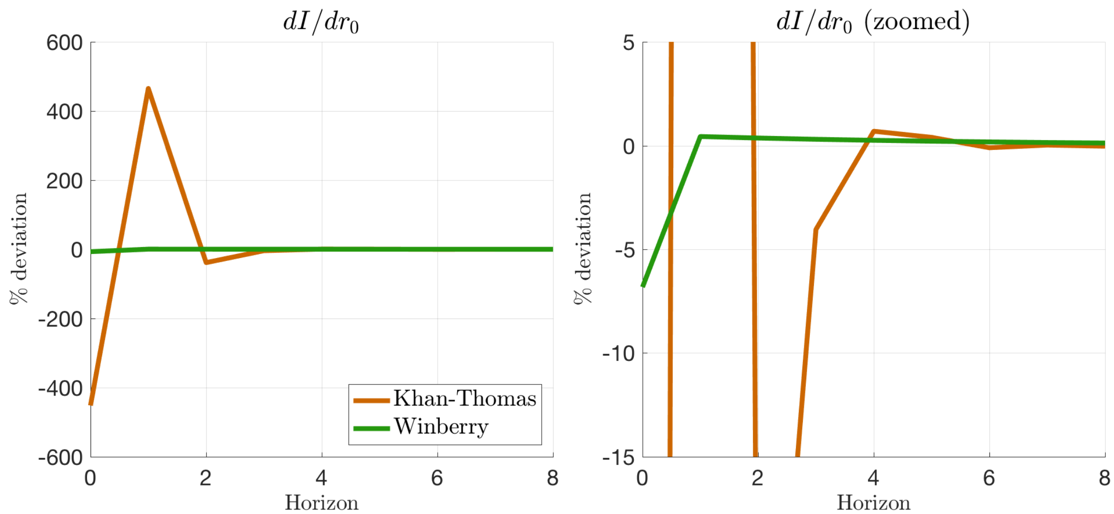

Figure 2: Response of partial equilibrium investment demand to a 100 basis point increase in real rates for one year in our replications of Khan & Thomas (2008) and Winberry (2018). We aggregate the sensitivities in Winberry (2018), which is a quarterly model, up to annual frequency.

In the structural model of Khan & Thomas (2008), investment is extremely price-sensitive. For example, a one per cent increase in interest rates for the current year lowers aggregate investment by almost 500 per cent, only to then see a similarly dramatic overshoot next period. In contrast, in the model of Winberry (2018), the average semi-elasticity is an order of magnitude lower, with total investment falling by a (still substantial) 7 per cent. Fur-

[7]To construct these plots, we consider the production sectors of each model in isolation, fix all prices at their deterministic steady-state value *except* for the current interest rate, and trace out the resulting behavior of aggregate investment.

thermore, the subsequent adjustment is much more drawn out, with firms only gradually rebuilding their capital stock.

The results in Figure 2 are entirely consistent with our aggregation theorem in Section 2.2: the price-elastic production block of Khan & Thomas (2008) aggregates, while the model in Winberry (2018) features substantial state dependence in the behavior of aggregate investment. The associated paths of aggregate prices are also consistent with our theory. In Khan & Thomas (2008), the response of real interest rates to an incremental productivity shock in a deep recession is only slightly smaller than the analogous response starting from a brisk expansion, while in Winberry (2018) the recession price drop is around 10 times larger than the boom drop.[8] Intuitively, investment demand in Khan & Thomas (2008) is highly elastic around the equilibrium price path of a standard neoclassical model without any fixed adjustment costs, so price and quantity impulse responses barely respond to changes in the cross-sectional distribution of capital.

Durables. Just like the formal aggregation result in Proposition 1, the more general equilibrium characterization offered in Proposition 3 also extends with little change to rich structural models of durables consumption, with details provided in Appendix A.3.

Our results again rationalize findings in previous work: In Berger & Vavra (2015) and McKay & Wieland (2019), the cross-sectional distribution of households over durable holdings (and so the implied adjustment hazards) are shown to quantitatively matter for the response of aggregate consumption to macroeconomic stimulus. The reason is simple: Standard durable consumption models allow for substantial curvature in household preferences over fast-depreciating durable consumption goods. Durables consumption demand is thus not highly elastic around the equilibrium price path implied by the corresponding representative-household economy, and so the consumption block does not aggregate.

3 Empirical evidence

We now turn to empirical evidence for the price elasticities that we have identified as central to the strength of general equilibrium smoothing. In Section 3.1 we argue that disagreement in previous work was possible because standard calibration targets are largely silent on the

[8] In Winberry (2018), the initial drop in a recession is -0.22%, while in a boom it is only -0.02% (see his Figure 6). In our calculations for the model of Khan & Thomas (2008), the recession impact impulse response is around 0.6%, while the boom response is 0.68%.

price sensitivity of firm-level investment. We thus in Section 3.2 directly measure these elasticities using firm-level quasi-experiments.

3.1 Standard calibration targets

Most previous contributions to the heterogeneous-firm investment literature did not target any direct empirical evidence on price elasticities (Khan & Thomas, 2008, 2013; Zetlin-Jones & Shourideh, 2017; Winberry, 2018; Ottonello & Winberry, 2018). Instead, price elasticities and so the strength of general equilibrium smoothing were disciplined indirectly through other calibration targets. We only discuss two prominent sets of target moments here, and relegate further discussion to Appendix B.2.

Investment lumpiness. All work building on Caballero & Engel (1999) targets the lumpiness of investment at the individual firm level. Our review of Khan & Thomas (2008) and Winberry (2018) in Section 2.3 reveals that lumpiness alone is simply not informative about price elasticities: Two models, consistent with the same empirically documented degree of firm-level investment lumpiness, can feature vastly different price elasticities and so differ in the extent of general equilibrium smoothing. This observation is not surprising given the results of our simple model in Section 2.2. Any given $\xi > 0$ – and so any given average lumpiness of investment – is consistent with either strong or weak general equilibrium smoothing, which instead depends on the curvature of firm revenue functions.

Aggregate prices. Winberry (2018) additionally targets empirical evidence on the conditional cyclicality of aggregate real interest rates. He finds that real interest rates move relatively little following identified aggregate technology shocks, and matches this response through sharply downward-sloping firm-level investment demand, thus indirectly breaking the strong general equilibrium smoothing of Khan & Thomas (2008). This indirect price-based inference approach, however, has two important limitations relative to our direct measurement in Section 3.2.

First, there is substantial uncertainty surrounding the measured target moment itself. Exogenous innovations to aggregate technology are notoriously difficult to identify in general (Ramey, 2016), and the response of real rates in particular is uncertain, not least because of likely interactions with the conduct of monetary policy (Görtz et al., 2019). With small negative real rate responses favoring downward-sloping investment demand, but small positive responses fully consistent with strong general equilibrium smoothing, the power of aggregate data to inform the aggregation debate is clearly limited. Second, even if perfectly measured, movements in real rates are at best only indirectly informative about the slope of investment demand. Intuitively, as a price, the equilibrium real interest rate reflects not only firm investment demand, but also household savings decisions as well as the elasticity of capital goods supply. In Appendix B.2 we formally show that an arbitrary interest rate cyclicality is in principle consistent with large price elasticities and so strong general equilibrium smoothing.

3.2 Direct measurement

Partial equilibrium price elasticities of investment demand should in theory be identifiable from firm-level (quasi-)experimental variation. For example, comparing investment behavior across two groups of firms – a treatment group that receives an interest rate subsidy, and a control group that does not – would allow researchers to pin down the interest rate elasticity of investment, $\partial\mathbf{I}/\partial\mathbf{r}$. Similar arguments would apply for the cost of capital, sales prices, and wage costs, thus giving direct discipline on the general equilibrium feedback map $\mathcal{H}$. Unfortunately, no such wealth of ideal firm-level quasi-experiments is available.

Extant evidence on firm-level investment behavior is nevertheless easily rich enough to discriminate between the extremes depicted in Figure 2. Our argument in this section proceeds in two steps. First, we show that evidence on a *particular* price elasticity – in our case the response of investment to tax-induced changes in the effective cost of capital – puts tight discipline on the larger set of price elasticities that we have identified as central to general equilibrium smoothing. Second, we review and extend estimates of the firm-level investment response to bonus depreciation tax stimulus, and establish that they provide a sharp and informative *upper bound* on investment price elasticities.

3.2.1 Linking elasticities

Our argument builds on the rich investment model of Section 2.1. In this model, a standard tax subsidy on firm investment expenditure directly affects the effective price of capital: $q_t = (1-\tau_t^i)\tilde{q}_t$, where $\tilde{q}_t$ is the actual price of capital goods, and τ_t^i is the subsidy.

Suppose for now that we could estimate the partial equilibrium elasticity of firm j's investment today with respect to a one-off change in this cost of capital – that is, $\partial i_{jt}/\partial\tau_t^i$ and so $\partial i_{jt}/\partial q_t$. Since changes in real rates also just change the effective cost of capital, we would intuitively expect $\partial i_{jt}/\partial q_t$ to be informative about $\partial i_{jt}/\partial r_t$. Similarly, since changes in expected wages and sales prices simply shift the return on investment as dictated by the firm revenue function $\pi_j = \pi(k_{-1j}, p^I, w, z \times e_j)$, we would expect $\partial i_{jt}/\partial q_t$ in conjunction with revenue function estimates to be informative about $\frac{\partial i_{jt}}{\partial w_{t+1}}$ and $\frac{\partial i_{jt}}{\partial p^I_{t+1}}$. Proposition 4 makes this argument precise in a simplified investment model.

Proposition 4. *Consider a variant of the firm problem (1), simplified to have fixed aggregate prices and productivity, no idiosyncratic productivity risk, no financial frictions, and a differentiable adjustment cost function, with $\phi_{k'}(k, k') = 0$ for $k' = k$. Let*

$$\beta_j \equiv \frac{\partial \log(i_{jt})}{\partial q_t}$$

denote the response of investment of firm j at time t to an unexpected one-off change in the post-tax price of capital q at time t. Then price elasticities satisfy

$$\begin{aligned}
\frac{\partial \log(i_{jt})}{\partial r_t} &= \beta_j \times \frac{1}{1+\bar{r}} \\
\frac{\partial \log(i_{jt})}{\partial p^I_{t+1}} &= -\beta_j \times \frac{\partial \pi(k_{-1}, p^I, w, z)}{\partial p^I} \times \frac{1}{1+\bar{r}} \\
\frac{\partial \log(i_{jt})}{\partial w^I_{t+1}} &= -\beta_j \times \frac{\partial \pi(k_{-1}, p^I, w, z)}{\partial w} \times \frac{1}{1+\bar{r}}
\end{aligned}$$

In Appendix B.3 we show numerically that, even in much richer models with idiosyncratic productivity risk and various frictions to capital accumulation, price elasticities remain inextricably linked. Identifying one is thus highly informative for all of them.[9]

3.2.2 Bonus depreciation

We use empirical evidence on the effects of temporary investment tax incentives to learn about partial equilibrium investment price elasticities. Our analysis builds closely on Zwick & Mahon (2017), who study the bonus depreciation stimulus policies of 2001-2004 and 2008-2010. In both episodes, firms were allowed to use an accelerated schedule to deduct the cost of investment purchases from taxable income – in other words, bonus alters the timing of deductions, but not their overall amount. Accelerated schedules thus promise to stimulate investment through discounting effects (for all firms) and present-day cash flow effects (for financially constrained firms).

[9] Proposition 4 is – like most of our subsequent discussion – restricted to contemporaneous price elasticities. It is, however, straightforward to see that the *dynamic* price elasticities entering the full intertemporal adjustment matrix $\mathcal{H}$ are similarly linked across different prices.

The extent to which a bonus depreciation policy stimulates a given firm j's investment depends sensitively on the duration of firm j's capital. For firms with short-lived capital goods, depreciation is fast anyway, so the ability to use an accelerated schedule has little value; conversely, for long-lived capital goods, the implied drop in the effective cost of capital can be substantial. Zwick & Mahon (2017) build on this logic to construct a firm-specific measure of exposure to bonus depreciation stimulus, b_{jt}, where b_{jt} is the firm-specific present value of implied tax savings for every dollar of additional investment. They then run regressions of the form

$$\log(i_{jt}) = \alpha_j + \delta_t + \beta_{ZM} \times b_{jt} + \text{controls} + \text{error} \tag{10}$$

where α_j is a firm-specific fixed effect, δ_t is a time fixed effect, and the regression is run weighted by investment shares. Our main result in this section is that the estimand β_{ZM} is robustly informative about the price elasticities that we have identified as central to general equilibrium smoothing.[10]

Interpretation. We interpret the regression (10) and in particular the estimand β_{ZM} through the lens of our structural model of firm investment (1). To map empirical design into model, we allow firms to differ in their capital depreciation rate δ_j. In this expanded model, a bonus depreciation policy for investment at time 0 is simply a firm-specific path $\{\tau_{jt}^b\}_{t=0}^{\infty}$, where firm j's investment i_{j0} at time 0 generates current and future cash receipts of τ_{jt}^b for $t = 0, 1, \ldots$.[11] Conveniently, in the absence of financial frictions, such a policy is isomorphic to a one-time, firm-specific investment tax subsidy.

Lemma 1. *Consider a variant of the firm problem (1) with heterogeneous depreciation rates $\{\delta_j\}$ and without financial frictions. Then the response paths of all aggregate prices and quantities in response to an unexpected bonus depreciation shock with firm-specific schedules $\{\tau_{jt}^b\}_{t=0}^{\infty}$ is identical to aggregate response paths after a period-0 firm-specific investment*

[10] We throughout focus on the static regression coefficient β_{ZM}, which is informative about static price elasticities and so allows us to distinguish between the two extremes depicted in Figure 2. We discuss dynamic adjustment patterns in Appendix B.4.

[11] Since bonus depreciation stimulus in practice is implemented as a reduction in tax liabilities, this is only strictly speaking true for firms with positive profit tax liabilities. Since the benchmark results of Zwick & Mahon (2017) are similar to those in a sample with only taxable firms, we for presentational simplicity ignore any profit tax considerations and parameterize a given bonus depreciation policy directly as such a path of rebates.

subsidy shock with

$$\tau^i_{j0} = \tau^b_{j0} + \sum_{t=1}^{\infty} \zeta^t \left(\prod_{q=1}^{\infty} \frac{1}{1 + r_{q-1}} \right) \tau^b_{jt} \tag{11}$$

Now suppose an econometrician were to run a simple model analogue of the empirical regression (10) of Zwick & Mahon at $t = 0$:

$$\widehat{\log(i_{jt})} = \alpha + \beta_{ZM} \times \tau^i_{jt} + \epsilon_{jt} \tag{12}$$

The following result characterizes the estimand β_{ZM}.

Proposition 5. *Consider a variant of the firm problem (1) with heterogeneous depreciation rates $\{\delta_j\}$ and without financial frictions. Let*

$$\beta \equiv \int_{s \in \mathcal{S}} \frac{\partial \log(i_t(s))}{\partial \tau^i_t} d\tilde{\mu}(s)$$

where $s = (e, k, \delta)$ is the firm state, and $\tilde{\mu}$ is the investment rate-weighted distribution of firms over their state space. Then, to first order, the estimand β_{ZM} in (12) satisfies

$$\beta_{ZM} \ = \ \beta \ + \ \underbrace{\frac{Cov_{\tilde{\mu}(s)}\left(\left(\frac{\partial \log(i_t(s))}{\partial \tau^i_t} - \beta\right) \tau^i_t(s), \tau^i_t(s)\right)}{Var_{\tilde{\mu}(s)}(\tau^i_t(s))}}_{\textit{selection effect}} + \underbrace{\frac{Cov_{\tilde{\mu}(s)}\left(\frac{\partial \log(i_t(s))}{\partial \mathbf{p}} \hat{\mathbf{p}}, \tau^i_t(s)\right)}{Var_{\tilde{\mu}(s)}(\tau^i_t(s))}}_{\textit{heterogeneous GE exposure}} \tag{13}$$

If the size of bonus depreciation stimulus to a given firm j is not systematically related to either (i) that firm's responsiveness to stimulus or (ii) the firm's exposure to aggregate general equilibrium effects induced by the stimulus, then the regression estimand β_{ZM} is exactly the desired partial equilibrium elasticity of aggregate investment with respect to changes in the cost of capital. Intuitively, since (12) leverages cross-sectional identifying information, it is of limited use to construct aggregate counterfactuals (Wolf, 2019), but is ideal to learn about partial equilibrium firm-side price elasticities and so the strength of general equilibrium smoothing.

For a first back-of-the-envelope calculation, we assume that (i) the empirical regression of Zwick & Mahon indeed perfectly maps into our model regression (12), and (ii) that all investment price elasticities are uniform across firms, ensuring that the selection and general equilibrium exposure terms are zero. The benchmark point estimate of Zwick & Mahon then indicates a partial equilibrium semi-elasticity of aggregate investment with respect to cost of capital changes of 7.2, and so – under the assumptions of Proposition 4 – an interest rate elasticity slightly smaller than that.[12] At face value, this back-of-the-envelope interpretation of the estimand in Zwick & Mahon provides strong evidence against the large elasticities of Khan & Thomas (2008), and is quite consistent with the point estimate of the interest rate elasticity in Winberry (2018).

Robustness & indirect inference. Our simple back-of-the-envelope interpretation of β_{ZM} is likely to be somewhat inaccurate, for at least three reasons. First, and most importantly, standard models of investment imply that firms with longer-lived capital goods are more responsive to changes in the cost of capital (see (3)), so the selection effect is likely to be positive. Similarly, if there is some general equilibrium crowding-out, the general equilibrium exposure term will be negative. Second, if some firms face financial constraints, then the present value of tax reductions ceases to be a sufficient statistic characterizing the overall bonus depreciation policy, and so we cannot interpret bonus depreciation policies as a simple change in the cost of capital. Third, the mapping from the empirical regression (10) to our model regression (12) is imperfect; in particular, the construction of b_{jt} in Zwick & Mahon (2017) is slightly different from the theoretical argument in (11).

Our preferred solution to all these concerns is simple indirect inference: Rather than directly interpreting the Zwick & Mahon estimand as the desired elasticity β, we simply replicate their estimand in a candidate structural model – inclusive of heterogeneity in depreciation rates and financial frictions –, choose a model parameterization to match the empirical estimate, and back out the implied true model-based tax elasticity β (as well as all other model-implied price elasticities). In Appendix B.4 we show how to do so in rich heterogeneous-firm models of investment, and in Section 4 we put this procedure to work and add the Zwick & Mahon moment to a set of other, more standard calibration targets of the extant heterogeneous-firm literature.

By and large, the results from our indirect inference approach suggest that the simple back-of-the-envelope calculation is a quite reliable guide, with our preferred estimate of the partial equilibrium interest rate elasticity of investment around 5 per cent. This number is somewhat below the naive interpretation of the benchmark estimates in Zwick & Mahon, largely reflecting the effects of selection: Firms with long-lived capital goods are also more exposed to the stimulus policy, so the selection bias is positive. Our back-of-the-envelope estimate – while orders of magnitude smaller than the large elasticities found in models with

[12] To arrive at this number, take the benchmark Zwick-Mahon point estimate of 3.7, and scale by the investment tax rate to arrive at $d\log(i_t)/dq_t$.

strong general equilibrium smoothing – is thus best viewed as an informative *upper bound.*

Dynamic adjustment. The regression (10) is annual and static. To circumvent problems of time aggregation and to learn about dynamic investment adjustment patterns in response to tax subsidies, we in Appendix B.4 extend the analysis of Zwick & Mahon (2017) to a quarterly Computstat sample, and then estimate a sequence of dynamic regressions to recover the full intertemporal adjustment pattern. We find, first, that small price elasticities are also a feature of this alternative high-frequency dataset, and second, that the data are inconsistent with the strong reversal dictated by high-elasticity neoclassical investment models, as displayed for example in the left panel of Figure 2.

4 Illustration: state dependence in policy analysis

This section illustrates that, in structural general equilibrium models consistent with empirical evidence on both (i) firm-level real and financial frictions to investment and (ii) the price elasticity of investment, the behavior of aggregate investment is decisively shaped by the cross-sectional distribution of firm capital holdings. Section 4.1 shows that the lumpiness of investment can dampen the effectiveness of monetary policy in classical TFP recessions, and Section 4.2 finds that the usual investment crowding-out effects associated with expansions in government spending are weakened or in fact even reversed in financial recessions. The computational algorithm used for all exercises in this section is described in Appendix E.

4.1 Monetary policy and lumpy investment

We embed a rich heterogeneous-firm block with lumpy firm investment into an otherwise standard medium-scale New Keynesian DSGE model. Details on model set-up and calibration are relegated to Appendix D.1 and Appendix D.2; for the purposes of our analysis here, it suffices to note that the heterogeneous-firm block is calibrated to be jointly consistent with firm-level investment lumpiness and the evidence on investment price elasticities reviewed in Section 3.2. We use this model to study the response of aggregate investment to expansionary monetary policy shocks over the business cycle, as a function of the underlying cross-sectional distribution of capital. Results are displayed in Figure 3.

The left panel shows how the impulse response of investment to an expansionary monetary policy shock varies over the cycle. We normalize the size of the shock so that, at the deterministic steady state, it increases impact investment by 1 per cent relative to steady

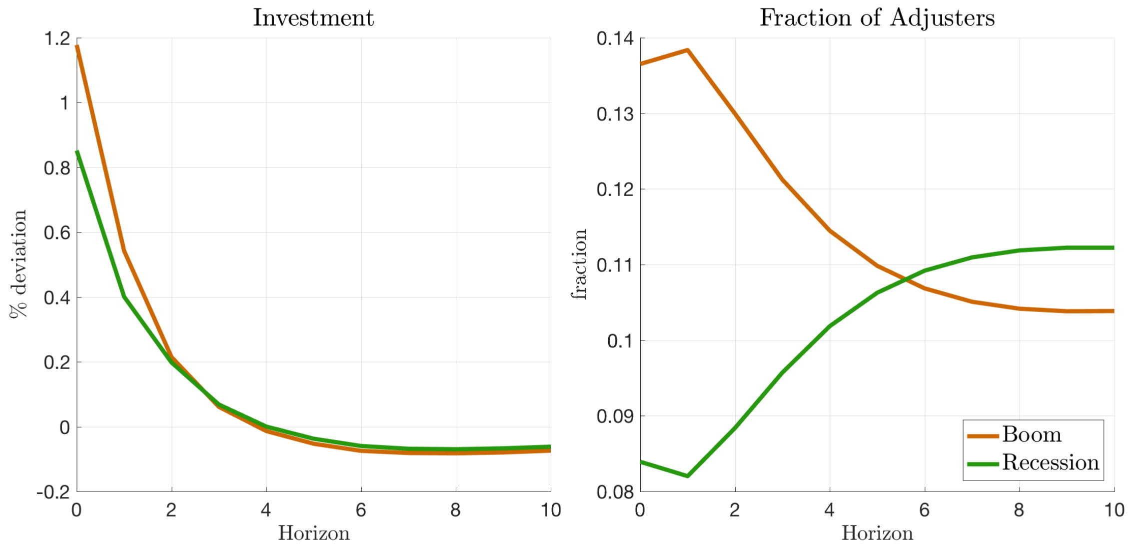

Figure 3: Response of aggregate investment to a monetary policy shock, normalized to increase investment by 1 per cent at the deterministic steady state. For the recession (boom) impulse response, we combine the monetary policy shock with a large contractionary (expansionary) TFP shock, lowering (increasing) output on impact by 5 per cent. The right panel shows the fraction of adjusting firms along the transition paths. A model period corresponds to one year.

state. For the displayed recession impulse response, we then complement this baseline expansionary monetary policy shock with a contractionary TFP shock, where the TFP shock is scaled to lower output on impact by 5 per cent on impact. For the boom case, we analogously consider an expansionary TFP shock that boosts output by 5 per cent. Strikingly, the elasticity of investment with respect to monetary policy shocks is strongly procyclical, with the impact stimulus to investment almost 40 per cent larger in a strong boom than in a similarly deep recession. The economic mechanism underlying this asymmetry is of course not novel – it is related to time variation in the extensive margin of capital adjustment, as analyzed in Caballero & Engel (1999) and as depicted here in the right panel. In a deep recession, the gradual depreciation of capital means that firms are on average close to their desired target capital, so they are unwilling to pay the fixed cost required for capital adjustment. With fewer firms adjusting, the response to any incremental shock – the expansionary monetary policy shock considered here – is dampened. Conversely, in a boom, more firms are adjusting anyway, and so the economy becomes more responsive to any incremental shock. The contribution of this paper is simply to show that, with data-consistent dampened price elasticities, this time variation in extensive margin adjustment probabilities survives to shape general equilibrium outcomes, in line with the logic presented in Section 2.

Similar to our results here, Winberry (2018) finds state dependence in the response of investment to tax credits. We emphasize, however, that our underlying structural model – unlike his – does not feature household habit formation. Habit formation may be necessary to match the behavior of real interest rates over the business cycle, but as we show in Section 2 and Appendix B.2, it is largely orthogonal to the question of (near-)aggregation, or the lack thereof. All that is needed to break aggregation are dampened price elasticities, and those are ensured both here and in Winberry (2018) through strong firm-level real adjustment costs. Finally, it is interesting to note that our overall conclusions receive some indirect empirical support. Notably, Tenreyro & Thwaites (2016) document substantial state dependence in the response of macroeconomic aggregates to monetary policy shocks; our analysis shows that the time-varying responsiveness of investment to aggregate shocks is a promising avenue to rationalize their findings.

4.2 Fiscal policy and financial frictions

For our second application, we begin with the same medium-scale DSGE model as before, but now allow for real *and* financial frictions in the heterogeneous-firm production block. Consistent with the empirical evidence in Lian & Ma (2018), we assume that firms face an earnings-based borrowing constraint. As before, details on the model outline and calibration are relegated to Appendix D.1 and Appendix D.3; we only emphasize that our benchmark calibration is consistent with the evidence of Zwick & Mahon on low price elasticities, and features a financially constrained fringe of firms with highly cash flow-responsive investment. In this environment, we study the response of investment to a fiscal spending expansion, first in normal times and then in a financial recession – a sudden tightening of the earnings-based borrowing constraint that leads to an output drop of around 2.5 per cent, consistent with the financial shock in Khan & Thomas (2013). The size of the transitory fiscal expansion is normalized to increase output on impact by 1 per cent.

The left panel shows the response of investment to the fiscal expansion in normal times (at the deterministic steady state) and when the economy was simultaneously hit by a tightening of the earnings-based borrowing constraint. In normal times, investment crowding-out is relatively limited. This is so because, even in normal times, there is a fringe of financially constrained and thus cash flow-dependent firms. A fiscal policy expansion pushes up real interest rates, crowding out investment, but at the same time boosts cash flow, crowding-in investment. In normal times these effects roughly offset, leading to a small investment

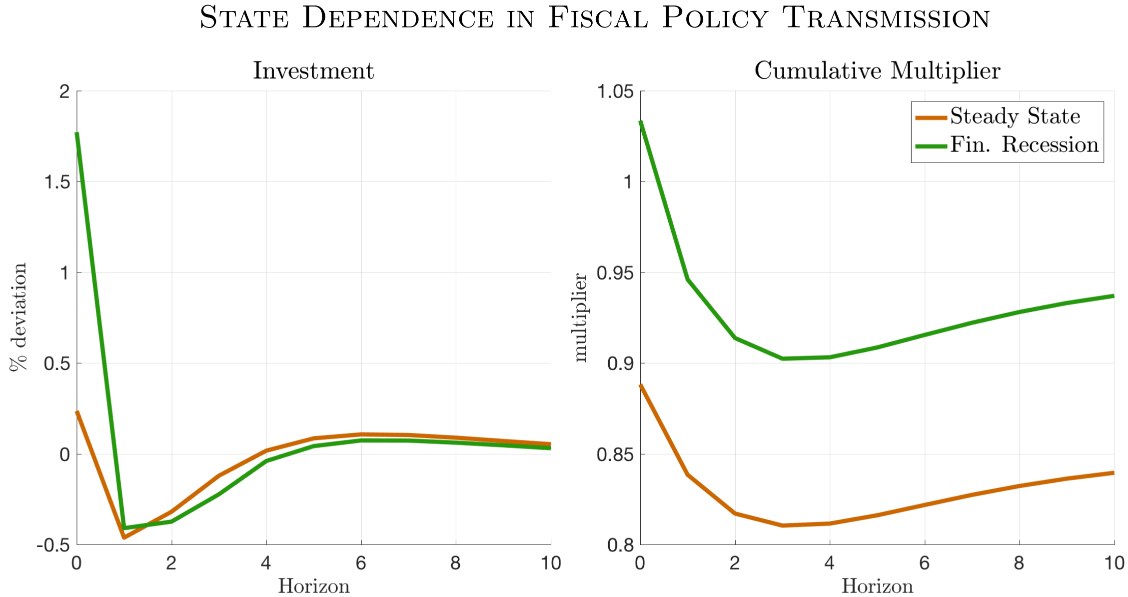

Figure 4: Response of aggregate investment to a fiscal policy shock, normalized to increase output by 1 per cent at the deterministic steady state. For the recession impulse response, we combine the fiscal policy shock with a sudden tightening in the earnings-based borrowing constraint, lowering output on impact by 2.5 per cent. The right panel shows cumulative government spending multipliers, defined as in Ramey (2016). A model period corresponds to one year.

response. Matters are different in crisis times: By construction, the brief demand boom induced by the transitory increase in government spending occurs at the same time as the tightening in firm borrowing constraints. The tightening of borrowing constraints pushed up the average cash flow sensitivity of firms, so the increase in demand now leads to substantial general equilibrium crowding-*in*. Importantly, since investment of unconstrained firms is not particularly price-elastic, the expansion in investment among constrained firms is not mirrored by a corresponding drop in investment among the unconstrained.[13] The right panel then shows that this state dependence in the investment response translates into non-negligible state dependence in aggregate government spending multipliers, with cumulative multipliers around 15% higher in a financial recession than in normal times. Overall, these results on the investment response to fiscal policy shocks are consistent with the empirical evidence reviewed in Wolf (2019). He finds limited general equilibrium crowding-out of

[13]Those results are echoed in Khan & Thomas (2013): In their model, meaningful investment irreversibility lowers the semi-elasticity of aggregate investment with respect to a change in the real interest rate to around -30 per cent, somewhat above our estimates, but materially below Khan & Thomas (2008). In light of our theory it is thus not surprising that, in their model, a tightening of financial constraints is not offset by reallocation of production across firms and so has important aggregate effects.

investment expenditure, exactly as in our model with financial frictions.

5 Conclusion

We identify the interest rate elasticity of firm investment as a key determinant of the strength of general equilibrium smoothing and so the extent to which the cross-sectional distribution of firm capital holdings shapes the behavior of aggregate investment. Our main empirical finding is that micro data strongly reject the large elasticities required for the general equilibrium smoothing documented in much previous work; instead, in models consistent with our measured small price elasticities, the lumpiness of firm investment as well as the presence of firm-level financial frictions matter for the transmission of monetary and fiscal policy.

Our analysis leaves several important avenues for future research. First, we have identified two possible offsetting effects on state dependence in policy transmission. For example, an investment tax credit may be less effective in recessions due to lumpy investment dynamics, or it may be more effective because financial constraints are particularly tight. Our results simply indicate that both effects *will* matter even in general equilibrium, but are silent on which one is more potent. Second, the investment frictions discussed here are likely to meaningfully interact with the micro frictions emphasized in recent work on heterogeneous-household models. For example, with household consumption less sensitively tied to real interest rates (Kaplan et al., 2018), investment plays a larger role in monetary policy transmission, and so the state dependence documented here will become even more pronounced in the aggregate. We leave a serious quantification of such interaction for future work.

References

Ahn, S., Kaplan, G., Moll, B., Winberry, T., & Wolf, C. K. (2017). When Inequality Matters for Macro and Macro Matters for Inequality. *NBER Macroeconomics Annual*, *32*.

Auclert, A., Bardóczy, B., Rognlie, M., & Straub, L. (2019). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. Technical report, National Bureau of Economic Research.

Auclert, A., Rognlie, M., & Straub, L. (2018). The Intertemporal Keynesian Cross. Working Paper.

Bachmann, R., Caballero, R. J., & Engel, E. M. (2013). Aggregate implications of lumpy investment: new evidence and a dsge model. *American Economic Journal: Macroeconomics*, *5*(4), 29–67.

Berger, D. & Vavra, J. (2015). Consumption dynamics during recessions. *Econometrica*, *83*(1), 101–154.

Boppart, T., Krusell, P., & Mitman, K. (2018). Exploiting mit shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, *89*, 68–92.

Brunnermeier, M. K. (2001). *Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding*. Oxford University Press on Demand.

Caballero, R. J. & Engel, E. M. (1999). Explaining investment dynamics in us manufacturing: a generalized (s, s) approach. *Econometrica*, *67*(4), 783–826.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, *113*(1), 1–45.

Clementi, G. L. & Palazzo, D. (2016). On The Calibration of Competitive Industry Dynamics Models. Unpublished working paper.

Cooper, R. W. & Haltiwanger, J. (2006). On the Nature of Capital Adjustment Costs. *Review of Economic Studies*, *73*, 611–633.

Cummins, J. G., Hassett, K. A., Hubbard, R. G., Hall, R. E., & Caballero, R. J. (1994). A reconsideration of investment behavior using tax reforms as natural experiments. *Brookings papers on economic activity*, *1994*(2), 1–74.

David, J., Schmid, L., & Zeke, D. (2018). Risk-adjusted capital allocation and misallocation. *Available at SSRN 3255111*.

Eberly, J., Rebelo, S., & Vincent, N. (2012). What explains the lagged-investment effect? *Journal of Monetary Economics*, *59*(4), 370–380.

Erceg, C. J., Henderson, D. W., & Levin, A. T. (2000). Optimal Monetary Policy with Staggered Wage and Price Contract. *Journal of Monetary Economics*, *46*(2), 281–313.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., & Schorfheide, F. (2016). Solution and estimation methods for dsge models. In *Handbook of Macroeconomics*, volume 2 (pp. 527–724). Elsevier.

Foster, L. S., Grim, C. A., Haltiwanger, J., & Wolf, Z. (2017). Macro and micro dynamics of productivity: From devilish details to insights. Technical report, National Bureau of Economic Research.

Gavazza, A., Mongey, S., & Violante, G. L. (2018). Aggregate Recruiting Intensity. *American Economic Review*, *108*(8), 2088—2127.

Goolsbee, A. (1998). Investment tax incentives, prices, and the supply of capital goods. *The Quarterly Journal of Economics*, *113*(1), 121–148.

Gorman, W. M. (1961). On a class of preference fields. *Metroeconomica*, *13*(2), 53–56.

Görtz, C., Korobilis, D., Tsoukalas, J. D., & Zanetti, F. (2019). The effect of news shocks and monetary policy.

House, C. L. (2014). Fixed costs and long-lived investments. *Journal of Monetary Economics*, *68*, 86–100.

House, C. L., Mocanu, A.-M., & Shapiro, M. D. (2017). Stimulus effects of investment tax incentives: Production versus purchases. Technical report, National Bureau of Economic Research.

House, C. L. & Shapiro, M. D. (2008). Temporary investment tax incentives: Theory with evidence from bonus depreciation. *American Economic Review*, *98*(3), 737–68.

Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, *57*(2), 132–145.

Kaplan, G., Moll, B., & Violante, G. (2018). Monetary Policy according to HANK. *American Economic Review*, *108*(3), 697–743.

Khan, A. & Thomas, J. K. (2008). Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics. *Econometrica*, *76*(2), 395–436.

Khan, A. & Thomas, J. K. (2013). Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. *Journal of Political Economy*, *121*(6), 1055–1107.

Lian, C. & Ma, Y. (2018). Anatomy of corporate borrowing constraints. Working Paper.

McKay, A. & Wieland, J. F. (2019). Lumpy durable consumption demand and the limited ammunition of monetary policy.

Ottonello, P. & Winberry, T. (2018). Financial Heterogeneity and the Investment Channel of Monetary Policy. Unpublished working paper.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of Macroeconomics*, volume 2 chapter 2, (pp. 71–162). Elsevier.

Sims, C. A. (2000). Solving Linear Rational Expectations Models. Unpublished working paper.

Smets, F. & Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, *97*(3), 586–606.

Tenreyro, S. & Thwaites, G. (2016). Pushing on a string: Us monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics*, *8*(4), 43–74.

Thomas, J. K. (2002). Is lumpy investment relevant for the business cycle? *Journal of political Economy*, *110*(3), 508–534.

Werning, I. (2016). Incomplete Markets and Aggregate Demand. Unpublished working paper.

Winberry, T. (2018). Lumpy Investment, Business Cycles, and Stimulus Policy. Unpublished working paper.

Wolf, C. K. (2019). The missing intercept: A sufficient statistics approach to general equilibrium effects. Working Paper.

Zetlin-Jones, A. & Shourideh, A. (2017). External financing and the role of financial frictions over the business cycle: Measurement and theory. *Journal of Monetary Economics*, *92*, 1–15.

Zwick, E. & Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *The American Economic Review*, *57*(3), 391–414.

A Details on aggregation theory

This appendix provides further details on the structural models and the aggregation theory of Section 2. We first discuss our assumptions about non-production model closures in Appendix A.1, then offer additional details on the firm investment problem in Appendix A.2, and finally show in Appendix A.3 that our aggregation logic extends with little change to a model with durable consumption.

A.1 Rest of the economy

This section provides additional details on our high-level assumptions about general equilibrium model closure. We first give a particular structural interpretation to our baseline Assumption 1, and then show that our aggregation results also survive under materially richer model closures.

A.1.1 Interpreting Assumption 1

In Appendix D.1, we sketch the rich medium-scale DSGE model framework into which we embed our heterogeneous-firm production blocks of Section 4. Now suppose that, rather than closing the production blocks in Section 2.1 and Section 2.2 with Assumption 1, we would instead consider this richer structural framework as our model closure, and define an equilibrium as in Definition D.1. As it turns out, with some additional restrictions, this extremely rich alternative framing model is in fact consistent with the high-level Assumption 1.

Lemma A.1. *Suppose that labor disutility is linear, that the coefficient on output in then Taylor rule is* 0*, and that there are no aggregate capital adjustment costs. Then, to first order, the full structural model of Appendix D.1 satisfies Assumption 1. If prices and wages are flexible, then Assumption 1 is satisfied globally.*

We can thus, for purposes of our analysis in Section 2, replace the cumbersome equilibrium definition Definition D.1 with the much simpler Definition 1. All aggregation results can then be interpreted as applying to a concrete *class* of models – our benchmark model of Appendix D.1, restricted in line with the extra assumptions in Lemma A.1.

Example illustrations. The illustrative plots in Figure 1 of course require a particular model parameterization. For the firm block, we choose $\alpha = 1/3$, $\delta = 0.025$, and $\xi = 0.5$; the value of the returns-to-scale parameter ν is varied across experiments. For the rest of the economy, we consider the $\mathbf{p}(\bullet)$-function induced by the simplest possible RBC closure. The representative household has a discount factor $\beta = 0.99$, log consumption preferences, and linear labor disutility. The TFP shock follows an AR(1) process, with persistence 0.9.

In this model, we could construct plots analogous to Figure 1 for arbitrarily different investment decision rules for constrained firms. For example, a randomly selected fraction of firms could be forced to at most invest its cash flow each period. Results for such alternative experiments are qualitatively identical to those in Figure 1 and available upon request.

A.1.2 Richer model closures

The full model outlined in Appendix D.1 violates Assumption 1. As emphasized in the main text, however, we impose this assumption simply for presentational simplicity and to emphasize the parallels between the economic intuition from simple static models and our more general infinite-horizon framework. In the most general model of Appendix D.1, the rest of the economy simply puts more complicated restrictions on the path of prices faced by the heterogeneous-firm block. Concretely, we replace Assumption 1 by the following more general restriction:

Assumption 2. *Let $\mathbf{x} = (\mathbf{y}, \mathbf{i}, \boldsymbol{\ell})$. There exists a function $\mathbf{p} = \mathbf{p}(\mathbf{x})$, mapping sequences of aggregate output, investment and hours worked into sequences of aggregate prices $\mathbf{p}$ such that all actors in the (non-production) rest of the economy behave optimally if and only if $\mathbf{p} = \mathbf{p}(\mathbf{x})$.*

We then consider the following generalized equilibrium definition:

Definition A.1. *Given an initial state s_0 and an exogenous aggregate TFP path $\{z_t\}_{t=0}^{\infty}$, a perfect foresight transition path equilibrium is a sequence of aggregate quantities $\{x_t\}_{t=0}^{\infty} = \{y_t, i_t, \ell_t\}_{t=0}^{\infty}$ and prices $\{r_t, p_t^I, w_t, q_t\}_{t=0}^{\infty}$ such that:*

1. Firm Optimization. *Given prices and aggregate TFP, the paths of aggregate investment $\mathbf{i} = \mathbf{i}(\mathbf{p}; \mathbf{z})$, labor demand $\boldsymbol{\ell} = \boldsymbol{\ell}(\mathbf{p}; \mathbf{z})$ and production $\mathbf{y} = \mathbf{y}(\mathbf{p}; \mathbf{z})$ are consistent with optimal intermediate goods producer behavior.*

2. Rest of the Economy. *Aggregate prices satisfy $\mathbf{p} = \mathbf{p}(\mathbf{x})$.*

3. Aggregate Consistency. *All markets clear for all $t = 0, 1, 2, \ldots$:*

$$y_t \;=\; e_1' \cdot x_t$$

$$
\begin{aligned}
i_t &= e_2' \cdot x_t \\
\ell_t &= e_3' \cdot x_t
\end{aligned}
$$

where e_i is the ith 3-dimensional unit basis vector.

Lemma A.2 establishes that the general pricing function $\mathbf{p}(\bullet)$ is precisely rich enough to nest even the most general model in Appendix D.1.

Lemma A.2. *To first order, the full model of Appendix D.1 satisfies Assumption 2.*

Crucially, our aggregation logic is completely independent of whether the restrictions on prices from the rest of the economy take the simple form in Assumption 1 or the more general form in Assumption 2:

Corollary A.1. *Suppose that either $\nu = 1$ or $\bar{r} + \delta = 0$, and that the equilibrium in Definition A.1 for the special case with exogenous ξ exists and is unique. Then the equilibrium price paths $\mathbf{p}$ as well as all equilibrium aggregates – total consumption $\mathbf{c}$, total investment $\mathbf{i}$, total labor hired $\boldsymbol{\ell}$ and total output $\mathbf{y}$ – are independent of ξ, so the initial distribution of firms μ_0 over their idiosyncratic state space is irrelevant.*

A.2 Details on the firm problem

We characterize the firm problem recursively. An initial state is a distribution over the associated firm state space, and firm-side aggregates are defined by integrating with respect to the measure of firms over the idiosyncratic state space, as usual.

Recursive formulation. We assume that firm productivity follows a Markov chain on a set $\mathcal{E}$ with transition probabilities π^e. The state variables of an individual firm's problem are its idiosyncratic productivity $e \in \mathcal{E}$, capital holdings $k \in \mathcal{K} \subset \mathbb{R}^+$, and total debt $b \in \mathcal{B} \subset \mathbb{R}$. The beginning-of-period value of a firm with state (e, k, b) satisfies

$$v_t(e,k,b) = (1-\zeta)v_t^d(e,k,b) + \zeta v_t^l(e,k,b) \tag{A.1}$$

where v_t^d is the value of an exiting firm, and v_t^l is the value of a continuing firm. An exiting firm hires labor to maximize its current-period payoff, sells its remaining capital stock, and re-pays all outstanding debt:

$$v_t^d(e,k,b) = \max_{\ell} \left\{ p_t^I z_t e f(k,\ell) - w_t \ell \right\} + (1-\delta)q_t k - (1+r_{t-1})b + \phi(k,0) \tag{A.2}$$

All proceeds from production and capital sales are returned to firm owners. Dying firms are replaced by newborns with initial capital $k_0 \in \mathcal{K}$ and initial debt $b_0 \in \mathcal{B}$. Initial productivity $e \in \mathcal{E}$ is drawn from a measure π_0^e.

A continuing firm similarly hires labor to maximize its current-period production value, but then decides on new investment and debt issuance subject to real and financial constraints. Furthermore the firm decides whether or not to pay out any dividends; in making this decision, it discounts future payoffs at the risk-free return r_t. The full value function conditional on survival satisfies

$$v_t^l(e,k,b) = \max_{i,k',b',d} d + \frac{1}{1+r_t}\mathbb{E}_e\left[v_{t+1}(e',k',b')\right] \tag{A.3}$$

such that

$$\begin{aligned} d &= \max_{\ell}\left\{p_t^I z_t e f(k,\ell) - w_t \ell\right\} - (1+r_{t-1})b - q_t i - \phi(k',k) + b' \\ k' &= (1-\delta)k + i \\ b' &\leq \Gamma(q_t k, \pi_t(e,k)) \\ d &\geq \underline{d} \end{aligned}$$

where $\pi_t(e,k) \equiv \max_\ell \left\{p_t^I z_t e f(k,\ell) - w_t \ell\right\}$. With Cobb-Douglas production function $f(k,\ell) = (k^\alpha \ell^{1-\alpha})^\nu$ we get $\pi_t(e,k) = \pi_t^* e^{\frac{1}{1-(1-\alpha)\nu}} k^{\frac{\alpha\nu}{1-(1-\alpha)\nu}}$ where π_t^* depends only on prices (p_t^I, w_t) and technology z_t (as well as parameters).

Aggregating across firms. Firm optimization gives a set of policy functions $\{\ell_t, i_t, k_t', b_t', d_t\}$ on the firm state space $\mathcal{E}\times\mathcal{K}\times\mathcal{B}$. The distribution of firms over the state space $\mathcal{S} = \mathcal{E}\times\mathcal{K}\times\mathcal{B}$ is summarized in the measure μ on $(\mathcal{S}, \Sigma_s)$, where Σ_s is the Borel σ-algebra over $\mathcal{S}$. The distribution evolves as $\mu' = \Lambda_t(\mu)$, where the mapping Λ_t is characterized via

$$\mu(S) = \pi^d \int_{\mathcal{E}} \mathbf{1}_{[(k_0,b_0,e)\in S]} d\pi_0^e(e) + (1-\pi_d)\int_{\mathcal{S}}\int_{\mathcal{E}} \mathbf{1}_{[(k_t'(s),b_t'(s),e')\in S]} d\pi^e(e,e') d\mu(s) \tag{A.4}$$

At each point in time t we can aggregate across firms to obtain aggregate output y_t, labor ℓ_t, capital k_t, investment i_t, and dividends d_t.

We assume that, for any path of prices $\mathbf{p}$ and technology $\mathbf{z}$ faced by firms and for any initial tuple of productivity, capital holdings and debt, this problem has a unique solution. Then, given a path of prices $\mathbf{p}$ and technology $\mathbf{z}$ as well as an initial (steady-state) distribution of firms over productivities e_{-1}, capital holdings k_{-1} and debt b_{-1}, this problem induces unique paths of aggregate output supply $\mathbf{y}$, investment demand $\mathbf{i}$, and labor demand $\boldsymbol{\ell}$.

A.3 Durables

Our aggregation results extend without change to models of durable consumption. We here first sketch a general equilibrium model with durable consumption, then provide a simple aggregation result analogous to Proposition 1, and finally relate our results to previous work.

A.3.1 Model

We make the same assumptions on aggregate risk and use the same notational conventions as in our benchmark heterogeneous-firm model in Section 2.1. Our focus is now the consumption-savings problem of households, and as before we summarize the rest of the economy through a simple aggregate pricing relation.

Households. There is a unit continuum of households i. The problem of household i is to

$$\max_{\{c_{it},d_{it},\ell_{it},b_{it}\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_{it}^{1-\gamma}-1}{1-\gamma} + \chi_d \frac{d_{it}^{1-\zeta}-1}{1-\zeta} - \chi_\ell \frac{\ell_{it}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right\} \right] \tag{A.5}$$

subject to the budget constraint

$$c_{it} + d_{it} + b_{it} = w_t e_{it} \ell_{it} + (1+r_{t-1}) b_{it-1} + (1-\delta) d_{it-1}$$

and the liquid asset borrowing constraint

$$b_{it} \geq \underline{b} + \kappa_d (1-\delta) d_{it}$$

It is straightforward to show that this household problem admits a recursive characterization with three idiosyncratic state variables: productivity e, liquid wealth b_{-1}, and past durable holdings d_{-1}. The initial distribution of households over this idiosyncratic state space is part of the initial aggregate state s_0.

Rest of the economy. We assume that the rest of the economy can be characterized through a small set of aggregate relations.

Assumption 3. *Given the initial state s_0, there exist functions $\mathbf{y} = \mathbf{y}(\mathbf{p})$, $\mathbf{i} = \mathbf{i}(\mathbf{p})$ and $\boldsymbol{\ell}^d = \boldsymbol{\ell}^d(\mathbf{p})$, mapping sequences of prices $\mathbf{p}$ into sequences of output, investment and labor* *demand such that all actors in the (non-household) rest of the economy behave optimally if and only if* $\mathbf{y} = \mathbf{y}(\mathbf{p})$, $\mathbf{i} = \mathbf{i}(\mathbf{p})$ *and* $\boldsymbol{\ell}^d = \boldsymbol{\ell}^d(\mathbf{p})$.

It is straightforward to show that a large family of structural general equilibrium models satisfies Assumption 3. The argument is analogous to the proof of Lemma A.2, and details are provided in Wolf (2019).

A.3.2 Aggregation

In heterogeneous-household models with durable consumption, the cross-sectional distribution of durables holdings shapes consumption demand in partial equilibrium, but becomes irrelevant in general equilibrium if durables demand is sufficiently price-elastic.

Simple aggregation result. It is straightforward to obtain an aggregation result analogous to Proposition 1 and an equilibrium characterization analogous to Proposition 2. Since the arguments are largely identical to our investment analysis, we only sketch key steps here.

For the simple aggregation result, we consider a variant of (A.5) without idiosyncratic earnings risk ($e_{it} = \bar{e}$ for all i, t), with a standard natural borrowing limit, and with a fraction $\xi \in (0, 1)$ of households forced to keep their durables holdings fixed. The log-linearized first-order condition characterizing optimal durables consumption of adjusting households is

$$-\zeta[1 - \beta(1-\delta)]\hat{\tilde{d}}_{it} = \hat{\tilde{\lambda}}_{it} - \beta(1-\delta)\mathbb{E}_t\left[\hat{\tilde{\lambda}}_{it+1}\right] \tag{A.6}$$

where $\lambda_{it} = c_{it}^{-\gamma}$ is the marginal utility of consumption, and tildes denote logs. With either $\zeta = 0$ or $\beta(1-\delta) = 1$, durable holdings become infinitely sensitive to changes in the intertemporal profile of the marginal utility of consumption and so real interest rates, exactly as in our main results on investment price elasticities. Given our restrictions on the model closure in Assumption 3, the argument can proceed exactly as in the proof of Proposition 1.

For an equilibrium characterization analogous to Proposition 2, we consider a generic shock $\boldsymbol{\varepsilon}$ to (A.5) that affects the path of durable and non-durable consumption, but leaves household labor supply unchanged. Let $x_{it} \equiv c_{it} + d_{it} - (1-\delta)d_{it-1}$ denote total household consumption expenditure. Then, given Assumption 3, it is easy to show that

$$\hat{\mathbf{p}}_{s_0} = \hat{\mathbf{p}}_{\bar{s}} + \mathcal{H} \times [\hat{\mathbf{x}}_{s_0}(\mathbf{p}_{\bar{s}}; \boldsymbol{\varepsilon}) - \hat{\mathbf{x}}_{\bar{s}}(\mathbf{p}_{\bar{s}}; \boldsymbol{\varepsilon})]$$

where now $\mathcal{H}$ is a matrix of consumption and labor demand as well as output and labor supply elasticities. In particular, in our simplified economy with exogenous ξ, the general equilibrium adjustment map $\mathcal{H}$ becomes rank-deficient as either $\zeta \to 0$ or $\beta(1-\delta) \to 1$.

B Details on measurement

This appendix provides further details on our empirical analysis of investment price elasticities. The first two sections elaborate on previous work, with Appendix B.1 detailing our replication of Khan & Thomas (2008) and Winberry (2018), and Appendix B.2 discussing how various other estimation targets relate to investment price elasticities. Appendix B.3 elaborates further on our argument that a single price elasticity is likely to be informative about all of them. Finally, in Appendix B.4, we provide further details on our interpretation of the Zwick & Mahon estimand as informative about the tax price elasticity of investment.

B.1 Replications

To construct the investment rate sensitivity plots Figure 2, we consider the structural models of Khan & Thomas (2008) and Winberry (2018), and solve – at their preferred parameterization – for the deterministic steady state, using the methods outlined in Appendix E.1. Then, proceeding as in Appendix E.2, we feed a long deterministic price path into the overall heterogeneous-firm block, with wages at their steady state value throughout, and interest rates at steady state at all times except for $t = 0$, where they are elevated by 100 basis points. The resulting partial equilibrium time paths of aggregate investment are depicted in Figure 2. Using the same approach, we can also numerically characterize the full partial equilibrium elasticity *matrices* $d\log(I)/dr$. Results are displayed in Figure B.1.

(a) Khan & Thomas (2008) **(b)** Winberry (2018)

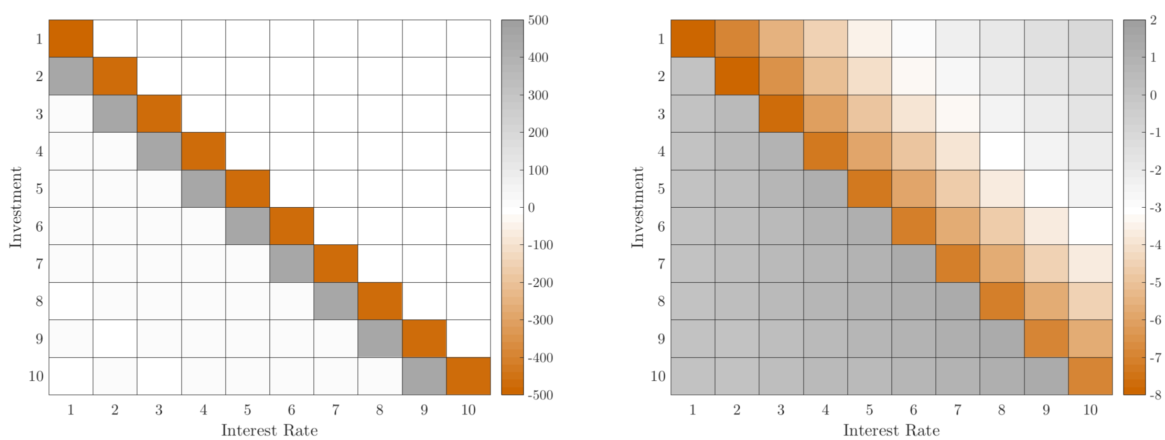

Figure B.1: Semi-elasticity of aggregate investment with respect to real interest rate by time horizon, our replications of Khan & Thomas (2008) and Winberry (2018). We aggregate the sensitivities in Winberry (2018), which is a quarterly model, up to annual frequency.

The first column of $d\log(I)/dr$ is simply the time path of the investment response displayed in Figure 2. Relative to this single response, the full elasticity matrices in Figure B.1 paint a richer picture of dynamic adjustment. In the model of Khan & Thomas (2008), the investment response follows a pronounced zig-zag pattern: investment drops in the period of the rate hike, and increases right after. In particular, there are no meaningful anticipation effects, nor is there delayed shock adjustment. As we show in Appendix B.4, this pattern of sharp adjustment and then reversal is inconsistent with micro investment data. In contrast, in the model of Winberry (2018), price elasticities are an order of magnitude smaller, and adjustment is more drawn out, consistent with the presence of significant adjustment costs.

B.2 Standard calibration targets

We here elaborate on the link between investment price elasticities and two popular calibration targets: the cyclicality of aggregate real interest rates and the dispersion of firm-level investment rates. While these moments are somewhat more directly linked to price elasticities than the lumpiness of investment itself, we nevertheless conclude that our direct evidence is likely to be more robust and so ultimately preferable.

Aggregate prices. Winberry (2018) rejects the aggregation result in Khan & Thomas (2008) on the grounds of its counterfactual implications for the conditional cyclicality of aggregate real interest rates. As reviewed in Section 3.1, this approach to model identification suffers from two potential defects: first, conditional interest rate responses are arguably not well-measured, and second, the theoretical link between prices and price elasticities can be tenuous. We here further discuss the second point.

Corollary A.1 reveals that our exact aggregation result is consistent with the presence of an arbitrarily inelastic aggregate capital goods producer (as for example familiar from the business-cycle New Keynesian literature). With such a model closure, the response of aggregate real interest rates to productivity shocks is muted, but the aggregation properties of the intermediate goods producer block of the economy are unaffected. Evidence on price responses alone is thus neither necessary nor sufficient to reject the (near-)aggregation documented in structural like that of Khan & Thomas (2008).

Investment rate dispersion. The link from price sensitivity to investment moments is crucially shaped by the firm-level productivity process. Intuitively, a high responsiveness to changes in prices – which govern the return profile of investment – also implies a high responsiveness to idiosyncratic productivity changes – which do the same. A dispersed idiosyncratic productivity distribution would thus, if firm investment were highly price-sensitive, map into a dispersed investment rate distribution. This investment rate distribution, however, is observable. In Khan & Thomas (2008) a high price sensitivity is consistent with a moderate investment rate dispersion only because the dispersion of idiosyncratic firm productivity is restricted to be extremely small.[14]

This insight suggests an alternative path to empirical identification of price sensitivities. Given direct estimates of idiosyncratic productivity processes, price sensitivities are indirectly disciplined through the investment rate distribution. In particular, given the empirically observed moderate dispersion in investment rates, a quite dispersed idiosyncratic productivity process – as suggested by most existing IO estimates – requires somewhat dampened price sensitivities. In fact it is encouraging to note that, given our estimated productivity process, the model estimation in Section 4 would imply relatively small price elasticities even when the Zwick & Mahon estimand is dropped from the list of targets. Of course, this alternative estimation is sensitively tied to auxiliary assumptions on the firm productivity process, and so the connection from estimation targets to the strength of general equilibrium smoothing would be tenuous at best. Identification through direct evidence on price elasticities in contrast is direct and so suffers from no such defects.

B.3 Linking price elasticities

To complement the theoretical result in Proposition 4, we here present a numerical illustration of how price elasticities in the general heterogeneous-firm models of Section 4 remain tied together. To construct Figure B.2 we begin with the benchmark parameterization of the structural model in Section 4.1, and then one-by-one vary either the upper bound of the fixed cost distribution $\bar{\xi}$ or the quadratic adjustment cost coefficient κ. For each resulting parameterization, we compute the semi-elasticity of aggregate investment with respect to a change in the cost of capital (τ_t^i, r_t) or the return on investment (w_{t+1}, p_{t+1}^I). Analogous results hold for other adjustment cost parameters (the divestment cost φ or the size of the penalty-free region a) as well as the model of Section 4.2; results are available upon request.

Visually, the different price elasticities appear to co-move, exactly as predicted by Proposition 4. Two features stand out. First, price elasticities are not hugely sensitive to variations in the fixed cost term $\bar{\xi}$, at least over the displayed range. To match the significantly damp-

[14] A mathematical formalization of this intuitive argument is available upon request.

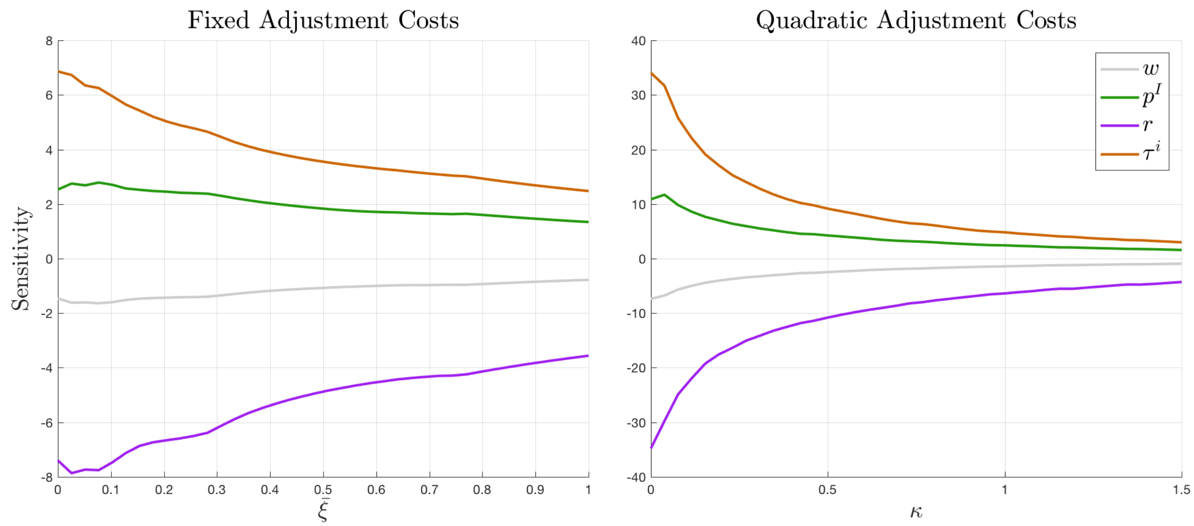

Figure B.2: Price elasticities of aggregate investment in the structural firm model of Section 4.1. Model is solved at the benchmark parameterization except for the fixed and quadratic adjustment cost parameters $\bar{\xi}$ and κ. The underlying grids for $\bar{\xi}$ and κ are uniform over the displayed ranges.

ened overall price elasticities, we instead need to choose a high κ. Second, even in this rich heterogeneous-firm model, the predictions of Proposition 4 remain not only qualitatively, but in fact quantitatively accurate. It is straightforward to show that, under the assumptions of the proposition, we have that $\frac{\partial \log(i_{jt})}{\partial r_t} \approx -\frac{\partial \log(i_{jt})}{\partial \tau_t^i}$, $\frac{\partial \log(i_{jt})}{\partial p_{t+1}^I} \approx -\frac{r+\delta}{1+r} \times \frac{1}{1-(1-\alpha)\nu} \times \frac{\partial \log(i_{jt})}{\partial \tau_t^i}$ as well as $\frac{\partial \log(i_{jt})}{\partial w_{t+1}} \approx \frac{r+\delta}{1+r} \times \frac{(1-\alpha)\nu}{1-(1-\alpha)\nu} \times \frac{\partial \log(i_{jt})}{\partial \tau_t^i}$. These relations continue to hold almost exactly in our richer heterogeneous-firm model, as is evident from Figure B.2.

B.4 Bonus depreciation

This section elaborates on our use of the bonus depreciation estimates of Zwick & Mahon (2017) to provide empirical discipline on the aggregation-relevant price elasticities of investment. We begin with further details on our indirect inference approach, then give some back-of-the-envelope intuition for the selection bias term, and finally discuss our additional empirical results, extending the baseline estimates of Zwick & Mahon (2017).

B.4.1 Details on indirect inference

Our indirect inference approach needs to deal with three challenges. First, the firm model in Section 4.2 features financial frictions, so we cannot rely on Lemma 1 to characterize the effects of bonus depreciation policy. Second, the policy experiments studied in Zwick

& Mahon (2017) were implemented in (financial) recessions and presumably had aggregate effects; we need to account for these. Third, we need to allow for heterogeneity in depreciation rates, giving the desired heterogeneous exposure of firms to the policy.

Financial Frictions. In a model with financial frictions, the Zwick-Mahon regression estimand is not invariant to the timing details of the implemented policy. Since a pure time-0 investment subsidy is a poor approximation to the implementation of actual bonus depreciation policies, the resulting biases could be significant. To alleviate those concerns, but maintain computational tractability, we consider a particular kind of bonus depreciation policy, designed to mimic the effects of the bonus depreciation stimulus observed in practice, yet at the same time keep the analysis simple. Consistent with the idea of bonus depreciation stimulus, we set $\tau_0^b > 0$ and $\tau_t^b \leq 0$ for $t = 1, 2, \ldots$, with $\sum_{t=0}^{\infty} \tau_t^b = 0$. Such a policy pulls cash payouts into the present, but provides no average stimulus beyond discounting effects. To maintain computational tractability we in fact set $\tau_t^b = 0$ for $t \geq 2$, so $\tau_1^b = -\tau_0^b$. We can then establish the following result.

Proposition B.1. *Let $\{r_t\}_{t=0}^{\infty}$ denote the equilibrium response path of real rates to an unexpected bonus depreciation shock with schedule $\{\tau_t^b\}_{t=0}^{\infty}$ such that $\tau_1^b = -\tau_0^b$ and $\tau_t^b = 0$ for $t \geq 2$. Then all impulse responses to the bonus depreciation shock are identical to impulse responses to a shock that sets period-0 investment subsidies to*

$$\tau_0^i = \tau_0^b + \frac{1}{1+r_0}\tau_1^b$$

and changes the period-0 borrowing constraint to

$$b' \leq \Gamma(q_0 k, \pi_0(e,k)) - \frac{1}{1+r_0}\tau_1^b i_0$$

The proposition shows that, in the presence of financial frictions, bonus depreciation policies have the additional effect of easing time-0 financial constraints (recall that $\tau_1^b < 0$), simply because firms are able to bring cash into the present. Computationally, the main appeal of our two-period implementation is that it allows us to avoid introducing an additional state variable to the firm problem.

Aggregate Effects. For the model in Section 4.1, we replicate the Zwick-Mahon regressions in a deep TFP recession, complemented with a one-year bonus depreciation stimulus of 7.8 cents, in line with the actual bonus depreciation episode of the Great Recession. For the model in Section 4.2, we do the same, but in a deep financial recession.

Heterogeneity in Depreciation Rates. Our benchmark models only feature a single common depreciation rate across all firms. To implement the Zwick-Mahon regressions, we thus add a measure 0 of firms with faster depreciation – 0.1 instead of 0.0667 –, compute the implied subsidy to each group of firms using standard MACRS depreciation schedules, and then use representative samples from both groups to replicate the regression estimands. Our results are relatively insensitive to the choice of depreciation rate for the control group, and in particular also go through almost without change if we instead considered equal (non-zero) masses of firms with different depreciation rates. We instead opt for our simpler benchmark with common depreciation rates to ease comparability to previous work. Detailed results for all alternative implementations are available upon request.

B.4.2 Selection bias

It is straightforward to show that, in standard models of investment, the price elasticity of investment is decreasing in the depreciation rate (House, 2014). Since the exposure to bonus depreciation $\tau_t^i(s)$ is mechanically decreasing in the depreciation rate, it follows that the selection effect bias is likely to be positive. In the language of treatment effects, the average treatment effect on the treated is larger than the treatment effect on the untreated, so the estimand β_{ZM} overstates the true average treatment effect β, as claimed in the text.

We can gauge the quantitative extent of the bias through a simple back-of-the-envelope calculation. Suppose for simplicity that the firm investment response to stimulus as well as the size of the stimulus itself depend *only* on the firm depreciation rate δ. It is then straightforward to show that, to first order,

$$\beta_{ZM} = \beta \times \left(1 + \frac{\beta_\delta}{\tau_\delta^i}\right) \tag{B.1}$$

where β_δ is the elasticity of the investment responsiveness with respect to the depreciation rate, and τ_δ^i is the corresponding elasticity of the stimulus. We approximate τ_δ^i from MACRS depreciation schedules, which roughly gives $\tau_\delta^i \approx -0.75$.[15] β_δ is harder to estimate. For simplicity, we just recover β_δ from our estimated structural models, giving $\beta_\delta \approx -0.4$. Thus,

[15] We use the depreciation schedules of $\delta = 0.1$ and $\delta = 0.067$ to construct a finite difference approximation.

according to (B.1), a benchmark estimate of β of around 7 per cent is mapped into a true average elasticity β^* of around 5 per cent, roughly consistent with the results of our full structural model estimation.

B.4.3 Additional empirical results

We have extended the analysis of Zwick & Mahon (2017) to trace out dynamic investment response patterns in a sample of Compustat firms, at quarterly frequency and ranging from 1993-2017. This additional empirical analysis is useful in several respects. First, by moving to a higher frequency, we assuage concerns about time aggregation. Second, our sample of Compustat firms is arguably less financially constrained than the larger sample of firms studied in Zwick & Mahon (2017). In the absence of financial frictions, the regression estimand is more closely tied to investment price elasticities (recall Proposition 4 and Lemma 1). Third, dynamic versions of the benchmark Zwick-Mahon regression are directly informative about the intertemporal capital adjustment pattern to tax stimulus. In particular, the dynamic regression estimands will offer a useful test of sharp reversal pattern characteristic of neoclassical models of investment (recall Figure 2). Finally, as a further robustness check, we add additional controls to firms' exposures to aggregate conditions throughout the sample.[16]

Results are displayed in Table 1. We run the regressions

$$\log(i_{jt+h}) = \alpha_j + \delta_t + \beta_h z_{n,t} + \text{controls} + \text{error}$$

where $z_{n,t}$ is the exposure variable for industry n and so firm j (or $\tau^i_{j,t}$ in the notation of the main text). Two results stand out. First, our estimated investment semi-elasticities are not too different from the benchmark Zwick-Mahon estimates; they are only somewhat smaller, consistent with the size dependence of tax elasticities documented in Zwick & Mahon (2017). Thus, if anything, our results affirm our main conclusions on dampened price elasticities and so weak general equilibrium smoothing. Second, the investment response gradually builds up over time. This pattern is at odds with the sharp reversal typical of infinite-elasticity neoclassical investment models, and more consistent with the lagged investment effect documented in Eberly et al. (2012). We leave a detailed matching of the full intertemporal path

[16]We add two further controls. First, we include an interaction term between a firm?s (industry) depreciation rate and aggregate GDP. Intuitively, we expect firms with higher average depreciation rates to be less price sensitive, and so potentially less cyclically exposed. Second, add a trend interaction, controlling for differential growth of high- vs. low- depreciation industries during that period. This is necessary since bonus depreciation expands during the sample.

of investment responses to future structural work.

Extension of Zwick & Mahon (2017)

Dependent Variable:	$\log(i_{j,t})$	$\log(i_{j,t+1})$	$\log(i_{j,t+2})$	$\log(i_{j,t+3})$	$\log(i_{j,t+4})$
$z_{n,t}$	1.64***	1.19***	0.78***	0.31	-0.12
	(0.28)	(0.28)	(0.29)	(0.29)	(0.30)
GDP Interaction	x	x	x	x	x
Trend Interaction	x	x	x	x	x
Firm & Time FEs	x	x	x	x	x
Observations	406,807	401,428	390,561	381,156	372,078
R-squared	0.85	0.85	0.85	0.86	0.86

Table 1: Note: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$. A period is quarter. Standard errors clustered at firm level.

C Proofs

C.1 Proof of Proposition 1

We show through guess-and-verify that, under the stated conditions, the equilibrium path of aggregate net output and so prices in an economy with $\xi = 0$ also is part of an equilibrium in the economy $\xi > 0$. Throughout, we use tildes and hats to denote *log* deviations from steady state.

Let $\pi_t^I = [1-(1-\alpha)\nu][(1-\alpha)\nu]^{\frac{(1-\alpha)\nu}{1-(1-\alpha)\nu}} \left(p_t^I z_t\right)^{\frac{1}{1-(1-\alpha)\nu}} w_t^{-\frac{(1-\alpha)\nu}{1-(1-\alpha)\nu}}$. Then for unconstrained firms we have

$$\hat{\tilde{k}}_{t+1,u} = \frac{1-(1-\alpha)\nu}{1-\nu} \times \left(\hat{\tilde{\pi}}_{t+1}^I - \frac{1+\bar{r}}{\bar{r}+\delta}\hat{r}_t\right)$$

Thus, for either $\nu = 1$ or $\bar{r}+\delta = 0$, the solution to the problem of unconstrained firms is not unique – they are indifferent about the scale of production. In line with our conjecture, we set

$$\hat{\tilde{k}}_{t,u} = \frac{1}{1-\xi} \times \left(\hat{\tilde{k}}_t(\bar{s}) - \xi\hat{\tilde{k}}_{t,c}\right)$$

where $\hat{\tilde{k}}_t(\bar{s})$ is capital in the economy with $\xi = 0$. Note that $\tilde{\hat{k}}_{t,c}$ is well-defined and finite.[17] Under the conjecture, the path of aggregate capital and so investment are thus identical to their paths in the economy with $\xi = 0$. Next note that labor hiring of adjusting firms satisfies

$$\hat{\tilde{w}}_t = \hat{\tilde{p}}_t^I + \alpha(\hat{\tilde{k}}_{t,u} - \hat{\tilde{\ell}}_{t,u})$$

But since labor hiring of constrained firms satisfies

$$\hat{\tilde{w}}_t = \hat{\tilde{p}}_t^I + \alpha(\hat{\tilde{k}}_{t,c} - \hat{\tilde{\ell}}_{t,c})$$

we can conclude immediately that *aggregate* labor hired is identical to the economy with $\xi = 0$. But then aggregate output, and so net aggregate output supply, are identical, and all markets clear, confirming the conjecture. □

[17] Actually it is 0 under our special assumptions. However, as emphasized throughout, our argument also applies for other (finitely price-sensitive) decision rules of constrained firms.

C.2 Proof of Proposition 2

The proof of (5) proceeds in three steps. First, we derive an implicit equilibrium characterization through a first-order expansion of (4). Second, we further simplify the resulting expression to arrive at an expression analogous to (5), but in net output space. Finally, we translate to price space to prove (5).

1. It is immediate that the pricing function $\mathbf{p}(\bullet)$ as well as the net output supply function $\tilde{\mathbf{y}}(\bullet)$ are differentiable, so to first order an equilibrium net output path $\tilde{\mathbf{y}}$ must satisfy

$$\hat{\tilde{\mathbf{y}}} = \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{z}} \times \hat{\mathbf{z}} + \mathcal{G} \times \hat{\tilde{\mathbf{y}}}$$

   Since the equilibrium is unique, we know that

$$\hat{\tilde{\mathbf{y}}} = \tilde{\mathcal{H}} \times \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{z}} \times \hat{\mathbf{z}}$$

   where $\tilde{\mathcal{H}}$ is a left inverse of $(I - \mathcal{G})$ (which exists by equilibrium existence, and is unique by equilibrium uniqueness).

2. Adding and substracting $\hat{\tilde{\mathbf{y}}}_{\bar{s}}$ on the right-hand side of (1), we obtain

$$\begin{aligned} \hat{\tilde{\mathbf{y}}} &= \hat{\tilde{\mathbf{y}}}_{\bar{s}} + \tilde{\mathcal{H}} \times \left(\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{z}} \times \hat{\mathbf{z}} - (I - \mathcal{G}) \hat{\tilde{\mathbf{y}}}_{\bar{s}} \right) \\ &= \hat{\tilde{\mathbf{y}}}_{\bar{s}} + \tilde{\mathcal{H}} \times \left(\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{z}} \times \hat{\mathbf{z}} + \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}} \times \hat{\mathbf{p}}_{\bar{s}} - \hat{\tilde{\mathbf{y}}}_{\bar{s}} \right) \\ &= \hat{\tilde{\mathbf{y}}}_{\bar{s}} + \tilde{\mathcal{H}} \times \left(\hat{\tilde{\mathbf{y}}}(\mathbf{p}_{\bar{s}}; \mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}} \right) \end{aligned}$$

   where the second line uses the definition of $\mathcal{G}$.

3. Since $\mathbf{p} = \mathbf{p}(\tilde{\mathbf{y}})$, we to first order have

$$\begin{aligned} \hat{\mathbf{p}} &= \frac{\partial \mathbf{p}}{\partial \tilde{\mathbf{y}}} \times \hat{\tilde{\mathbf{y}}} \\ &= \frac{\partial \mathbf{p}}{\partial \tilde{\mathbf{y}}} \times \left[\hat{\tilde{\mathbf{y}}}_{\bar{s}} + \tilde{\mathcal{H}} \times \left(\hat{\tilde{\mathbf{y}}}(\mathbf{p}_{\bar{s}}; \mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}} \right) \right] \\ &= \hat{\mathbf{p}}_{\bar{s}} + \mathcal{H} \times \left(\hat{\tilde{\mathbf{y}}}(\mathbf{p}_{\bar{s}}; \mathbf{z}) - \hat{\tilde{\mathbf{y}}}_{\bar{s}} \right) \end{aligned}$$

   as claimed.

It remains to show the null space result. We begin by studying the limit of the benchmark economy with $\xi = 0$ as $\nu \to 1$ (the argument for the case $\bar{r} + \delta \to 0$ is identical). Since the argument is to first order anyway, we study the linearized system of expectational stochastic difference equations characterizing the equilibrium in *gensys* form (Sims, 2000):

$$\Gamma_0 x_t = \Gamma_1 x_{t-1} + \Pi \eta_t + \Psi \varepsilon_t$$

where as usual x_t collects all model variables. Since the optimality conditions characterizing the solution to the firm problem are continuous in ν (including at the boundary $\nu = 1$), the matrices $(\Gamma_0, \Gamma_1, \Pi, \Psi)$ are continuous in ν. By assumption there exists a unique bounded solution to the system of expectational difference equations for all ν in a neighborhood $(\underline{\nu}, 1]$, so that solution must be continuous in ν. Thus $\lim_{\nu \to 1} \hat{\hat{\mathbf{y}}}_{\bar{s}}$ exists and is finite. Next let $\hat{\hat{\mathbf{y}}}_u(\mathbf{p}_{\bar{s}}; \mathbf{z})$ denote net output supply of unconstrained firms, and similarly with $\hat{\hat{\mathbf{y}}}_c(\mathbf{p}_{\bar{s}}; \mathbf{z})$ for constrained firms, so that total net output supply satisfies

$$\hat{\hat{\mathbf{y}}}(\mathbf{p}_{\bar{s}}; \mathbf{z}) = \xi \hat{\hat{\mathbf{y}}}_c(\mathbf{p}_{\bar{s}}; \mathbf{z}) + (1 - \xi) \hat{\hat{\mathbf{y}}}_u(\mathbf{p}_{\bar{s}}; \mathbf{z})$$

At $\nu = 1$ the solution to the *partial equilibrium* firm problem is not unique, but using the unique general equilibrium of the economy with $\xi = 0$ we have $\lim_{\nu \to 1} \hat{\hat{\mathbf{y}}}_u(\mathbf{p}_{\bar{s}}; \mathbf{z}) = \hat{\hat{\mathbf{y}}}_{\bar{s}}$. Finally $\lim_{\nu \to 1} \hat{\hat{\mathbf{y}}}_c(\mathbf{p}_{\bar{s}}; \mathbf{z})$ trivially exists and is finite, but generically is not equal to $\lim_{\nu \to 1} \hat{\hat{\mathbf{y}}}_u(\mathbf{p}_{\bar{s}}; \mathbf{z})$, establishing the desired conclusion. □

C.3 Proof of Proposition 3

Since the equilibrium in the general model is characterized by the fixed-point relation (4), the proof of (8) can proceed in the same three steps as the proof of the first part of Proposition 2. □

C.4 Proof of Proposition 4

For a time-varying path of prices, we can characterize the firm problem recursively as follows:

$$v_t(k) = \pi(k; w_t, p_t^I) - q_t(k' - (1 - \delta)k) - \phi(k, k') + \frac{1}{1 + r_t} V_{t+1}(k')$$

Optimal investment is then characterized by the single FOC

$$(1+r_t)[q_t+\phi_{k'}(k,k')]-q_t(1-\delta)+\phi_k(k',k'')=\frac{\partial}{\partial k}\pi(k';w_{t+1},p^I_{t+1})$$

We study the effects of one-off changes in $(q_t, r_t, w_{t+1}, p^I_{t+1})$ on investment at time t, with prices fixed at steady state in all future time periods. Since k is the single endogenous state of the firm problem, we know that $k' = k'(k, \bullet)$, and similarly $k'' = k'(k', \bullet)$. We can thus more concisely write

$$f(k,k';q_t,r_t,w_{t+1},p^I_{t+1})=0$$

where the steady-state assumption on prices ensures that no further time dependence is needed. Totally differentiating, we obtain that

$$[q_t+\phi_{k'}(k,k')]dr_t=f_{k'}(k,k';q_t,r_t,w_{t+1},p^I_{t+1})\times\frac{dk_{t+1}}{k}$$

and similarly

$$\begin{aligned}
(1+r)dq_t &= f_{k'}(k,k';_t,\tau^i_t,r_t,w_{t+1},p^I_{t+1})\times\frac{dk_{t+1}}{k}\\
\frac{\partial\pi(k';w,p^I)}{\partial p^I}dp^I_t &= f_{k'}(k,k';q_t,r_t,w_{t+1},p^I_{t+1})\times\frac{dk_{t+1}}{k}\\
\frac{\partial\pi(k';w,p^I)}{\partial w}dw_t &= f_{k'}(k,k';q_t,r_t,w_{t+1},p^I_{t+1})\times\frac{dk_{t+1}}{k}
\end{aligned}$$

But since we depart from steady-state prices and so steady-state capital, we know that initially $k' = k$. Since $\phi_{k'}(k, k) = 0$ by assumption, the statement follows. □

C.5 Proof of Lemma 1

We prove the statement in the case of no firm death, as the generalization to firm death only causes notational difficulties. Following Appendix A.2 we can characterize the firm problem recursively as

$$v_t(e,k,\iota_0)=\begin{cases}\max_{k'}\tau^b_t(k'-(1-\delta)k)+\pi_t(e,k)-(k'-(1-\delta)k)-\phi(k,k') & \\ +\frac{1}{1+r_t}\mathbb{E}_e\left[v_{t+1}(e',k',k'-(1-\delta)k)\right] & \text{if } t=0\\ \max_{k'}\tau^b_t\iota_0+\pi_t(e,k)-(k'-(1-\delta)k)-\phi(k,k') & \\ +\frac{1}{1+r_t}\mathbb{E}_e\left[v_{t+1}(e',k',\iota_0)\right] & \text{if } t>0\end{cases}$$

Pulling future cash returns out of the definition of the value function, we can equivalently write the problem as

$$v_t(e,k) = \begin{cases} \max_{k'} \left[\tau_0^b + \sum_{t=1}^{\infty} \left(\prod_{q=1}^{\infty} \frac{1}{1+r_{q-1}}\right) \tau_t^b\right] (k' - (1-\delta)k) + \pi_t(e,k) & \\ -(k' - (1-\delta)k) - \phi(k,k') + \frac{1}{1+r_t} \mathbb{E}_e \left[v_{t+1}(e',k')\right] & \text{if } t = 0 \\ \max_{k'} \pi_t(e,k) - (k' - (1-\delta)k) - \phi(k,k') + \frac{1}{1+r_t} \mathbb{E}_e \left[v_{t+1}(e',k')\right] & \text{if } t > 0 \end{cases}$$

and the desired conclusion is immediate. □

C.6 Proof of Proposition 5

Let tildes and hats denote *log* deviations from steady state. Then, to first order,

$$\hat{\tilde{i}}_t(s) = \frac{\partial \hat{\tilde{i}}_t(s)}{\partial \tau_t^i} \tau_t^i(s) + \frac{\partial \hat{\tilde{i}}_t(s)}{\partial \mathbf{p}} \hat{\mathbf{p}}$$

We thus have

$$\beta_{ZM} = \frac{\mathrm{Cov}_{\tilde{\mu}(s)}\left(\frac{\partial \log(i_t(s))}{\partial \tau_t^i} \tau_t^i(s), \tau_t^i(s)\right)}{\mathrm{Var}_{\tilde{\mu}(s)}(\tau_t^i(s))} + \frac{\mathrm{Cov}_{\tilde{\mu}(s)}\left(\frac{\partial \log(i_t(s))}{\partial \mathbf{p}} \hat{\mathbf{p}}, \tau_t^i(s)\right)}{\mathrm{Var}_{\tilde{\mu}(s)}(\tau_t^i(s))}$$

Simply adding and subtracting the definition of β gives (13). □

C.7 Proof of Lemma A.1

The proof is constructive. We are given a path of net output supply $\tilde{\mathbf{y}}$. Since the path of government spending is known from (D.11), we know the path of consumption. From (D.1), this gives us the path of marginal utility $\boldsymbol{\lambda}$. With (D.3) this gives the path of real rates. In the absence of aggregate capital adjustment costs, (D.10) implies that $\mathbf{q} = \mathbf{1}$. Next, if wages and prices are flexible, then also $\mathbf{p}^I = \mathbf{1}$; if additionally $\varphi = \infty$, then we immediately obtain $\mathbf{w}$ from (D.2). With sticky prices and wages, assuming that $\phi_y = 0$, and given $\mathbf{r}$, we can combine (D.2) and (D.14) to recover $\mathbf{r}^n$ and $\boldsymbol{\pi}$. Since to first order (D.7) is independent of $\mathbf{y}$, we obtain $\mathbf{p}^I$ from (D.6) - (D.7). Finally note that, if $\varphi = \infty$, then (D.4) is, to first order, independent of $\boldsymbol{\ell}$. We can thus combine (D.4) and (D.5) to get $\mathbf{w}$.

We have thus proven the existence of a function mapping $\tilde{\mathbf{y}}$ into $p = (r, p^I, w, q)$. In deriving this mapping, we have ensured that (D.1) - (D.5), (D.6) - (D.7), (D.11), (D.14), (D.10) and (D.15) all hold. (D.8) and (D.12) hold residually. By linear labor disutility, (D.17) holds. Finally (D.16) follows by Walras' law. Thus an equilibrium satisfies Definition D.1 if and only if it satisfies Definition 1, as claimed. □

C.8 Proof of Lemma A.2

The proof is exactly analogous to that of Lemma A.1. We are given $\mathbf{x} = (\mathbf{y}, \mathbf{i}, \boldsymbol{\ell})$. We thus know $\tilde{\mathbf{y}}$, and with the path of government spending known from (D.11), we know the path of consumption. From (D.1), this gives us the path of marginal utility $\boldsymbol{\lambda}$. With (D.3) this gives the path of real rates. Next, given the path of investment $\mathbf{i}$ and real rates $\mathbf{r}$, we can recover $\mathbf{q}$ from (D.10). Using $\mathbf{r}$ and $\mathbf{y}$, we can combine (D.2) and (D.14) to recover $\mathbf{r}^n$ and $\boldsymbol{\pi}$. We thus also obtain $\mathbf{p}^I$ from (D.6) - (D.7). Finally, given $\mathbf{c}$, $\boldsymbol{\ell}$ and $\boldsymbol{\pi}$, we can combine (D.4) and (D.5) to get $\mathbf{w}$.

We have thus proven the existence of a function mapping $\mathbf{x}$ into $p = (r, p^I, w, q)$. In deriving this mapping, we have ensured that (D.1) - (D.5), (D.6) - (D.7), (D.11), (D.14), (D.10) and (D.15) all hold. Finally market-clearing is ensured by the consistency requirements in Definition A.1. Thus an equilibrium satisfies Definition D.1 if and only if it satisfies Definition A.1, as claimed. □

C.9 Proof of Corollary A.1

This result is an immediate implication of the proof of Proposition 1. Under the conjecture

$$\hat{\hat{k}}_{t,u} = \frac{1}{1-\xi} \times \left(\hat{\hat{k}}_t(\bar{s}) - \xi \hat{\hat{k}}_{t,c}\right)$$

total output produced, investment demanded, and labor demanded are identical to the economy with $\xi = 0$. But then all of the extended consistency requirements of Definition A.1 still hold, so we are done. □

C.10 Proof of Proposition B.1

We can write the firm problem at time 0 as follows:

$$v_0(e, k, b) = (1 - \xi)v_0^d(e, k, b) + \xi v_0^l(e, k, b)$$

with exit value

$$v_0^d(e,k,b) = \pi_t(e,k) + (1-\delta)q_0 k - (1+r_{-1})b + \phi(k,0)$$

and full value conditional on survival

$$v_0^l(e,k,b) = \max_{i,k',b',d} d + \frac{1}{1+r_0}\mathbb{E}_e\left[v_1(e',k',b' - \frac{1}{1+r_0}\tau_1^b i)\right]$$

such that

$$\begin{aligned} d &= \pi_0(e,k) - (1+r_{-1})b - (1-\tau_0^b)q_0 i - \phi(k,k') + b' \\ k' &= (1-\delta)k + i \\ d &\geq \underline{d} \\ b' &\leq \Gamma(q_0 k, \pi_0(e,k)) \end{aligned}$$

Re-defining $\tilde{b}' = b' - \frac{1}{1+r_0}\tau_1^b$, the result follows. $\square$

D Details on applications

D.1 Rest of the economy

For the analysis in Section 4, we embed rich heterogeneous-firm production blocks into an otherwise standard medium-scale business-cycle model. In addition to intermediate goods producers, the model is populated by four other groups of agents: First, households consume, supply labor subject to standard sticky-wage frictions, and save in nominal bonds as well as real loans to the production block. Second, retailers purchase the homogenous good produced by the heterogeneous-firm block, costlessly differentiate it, and set prices subject to a sticky-price friction. Third, the common final consumption good is transformed into an investment good by a perfectly competitive capital goods producer. And fourth, government and monetary authority tax, spend, and set the rate on nominal bonds.

Households. There is a unit continuum $i \in [0, 1]$ of perfectly insured households with heterogenous labor varieties and habit formation in consumption. They provide their labor to a union, which sets prices for each type of labor and sells the labor to a competitive labor packer. Households use labor income and dividend rebates from the heterogeneous-firm production block as well as the retailer block to pay taxes, consume and save, either in nominal government-issued debt or real debt issued by the production sector. Given a wage and labor demand path, the problem of household i is thus

$$\max_{\{c_t(i), \ell_t(i), b_t(i), b_t^n(i)\}} \quad \sum_{t=0}^{\infty} \beta^t \left\{ \frac{(c_t(i) - hc_{t-1}(i))^{1-\gamma} - 1}{1-\gamma} - \chi \frac{\ell_t(i)^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right\}$$

such that

$$c_t(i) + b_t^n(i) + b_t(i) = w_t \ell_t(i) + \frac{1 + r_{t-1}^n}{1 + \pi_t} b_{t-1}^n(i) + (1 + r_t) b_{t-1}(i) + d_t + d_t^R + d_t^Q - \tau_t$$

where c denotes consumption, ℓ denotes labor, b^n denotes nominal bonds, b denotes real bonds, r is the real rate, r^n is the nominal rate, d denotes total dividends paid out by the heterogenous-firm production block, d^R denotes total retailer dividends, d^Q denotes total capital goods producer dividends, and τ is total tax payments. By perfect insurance, optimal savings and consumption decisions are summarized by the relations

$$\lambda_t \quad = \quad (c_t - hc_{t-1})^{-\gamma} - \beta h \left[(c_{t+1} - hc_t)^{-\gamma} \right] \tag{D.1}$$

$$\lambda_t = \beta \cdot [\frac{1+r_t^n}{1+\pi_{t+1}} \times \lambda_{t+1}] \tag{D.2}$$

$$\lambda_t = \beta \cdot [(1+r_{t+1}) \times \lambda_{t+1}] \tag{D.3}$$

where λ_t denotes the *common* stochastic discount factor of households. Following Erceg et al. (2000), wages are set by a union subject to wage adjustment costs; in our case, we take those to come in the form of Rotemberg adjustment costs. Standard algebra gives the optimal wage-setting relation

$$(1+\pi_t^w)\pi_t^w = \frac{\epsilon_w \chi \ell_t^{\frac{1}{\varphi}}}{\theta_w \lambda_t} + \frac{1-\epsilon_w}{\theta_w} w_t - \frac{1}{2}\frac{1-\epsilon_w}{\lambda_t}(\pi_t^w)^2 + \beta\left[\frac{\lambda_{t+1}}{\lambda_t}\pi_{t+1}^w(1+\pi_{t+1}^w)\frac{\ell_{t+1}}{\ell_t}\right] \tag{D.4}$$

where $\ell_t = \int_0^1 \ell_t(i)di$ is total labor and where wage inflation is defined as

$$1+\pi_t^w = \frac{w_t}{w_{t-1}}(1+\pi_t) \tag{D.5}$$

The household block of the economy is fully summarized by equations (D.1) - (D.5).

Retailers. A competitive final goods producer purchases retail goods $k \in [0,1]$ at price $p_t(k)$ and aggregates them into the final good, with production function[18]

$$y_t = \left[\int_0^1 y_t(k)^{\frac{\epsilon_p-1}{\epsilon_p}} dk\right]^{\frac{\epsilon_p}{\epsilon_p-1}}$$

The problem of the aggregator gives demand functions

$$y_t(k) = \left(\frac{p_t(k)}{p_t}\right)^{-\epsilon_p} y_t$$

and the price index

$$p_t = \left(\int_0^1 p_t(k)^{1-\epsilon_p} dk\right)^{\frac{1}{1-\epsilon_p}}$$

Retail goods are produced by a unit continuum of retailers $k \in [0,1]$. Retailers can costlessly turn one unit of the intermediate good produced by the heterogeneous-firm production block into one unit of their retail good, and for the retail good face the demand function derived

[18] Since in equilibrium there is no price dispersion, we can use the same y_t for the output of the heterogeneous-firm production block and the final output good.

from the problem of the aggregator. In setting their prices, they are subject to Rotemberg adjustment costs. The price-setting problem of a single retailer k is thus

$$\max_{\{p_{t+s}(k)\}} \mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s \lambda_{t+s} \left[p_{t+s}(k)^{1-\epsilon_p} p_{t+s}^{\epsilon_p - 1} y_{t+s} - mc_{t+s} p_{t+s}(k)^{-\epsilon_p} p_{t+s}^{\epsilon_p} y_{t+s} \right.\right.$$
$$\left.\left. - \frac{\theta_p}{2} \left[\frac{p_{t+s}(k)}{p_{t+s-1}(k)} - 1 \right]^2 p_{t+s}(k)^{1-\epsilon_p} p_{t+s}^{\epsilon_p - 1} y_{t+s} \right] \right]$$

where real marginal costs mc are given as

$$mc_t = p_t^I \tag{D.6}$$

Taking the FOC and using symmetry across firms, we get the non-linear NKPC

$$(1+\pi_t)\pi_t = \frac{1-\epsilon_p}{\theta_p} + \frac{\epsilon_p}{\theta_p} mc_t - \frac{1}{2}(1-\epsilon_p)\pi_t^2 + \beta \left[\frac{\lambda_{t+1}}{\lambda_t} \pi_{t+1}(1+\pi_{t+1}) \frac{y_{t+1}}{y_t} \right] \tag{D.7}$$

Total profits (in real terms) are given as

$$d_t^R = (1 - p_t^I) y_t \tag{D.8}$$

The retail block of the economy is fully summarized by equations (D.6) - (D.8).

Capital Goods Producer. The capital goods producer purchases the final output good, transforms it into capital, and sells capital at price $\tilde{q}_t$ to intermediate goods producers. Its problem is thus to

$$\max_{\{i_t\}} \quad \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \lambda_t i_t \left\{ \tilde{q}_t \left[1 - \frac{\kappa}{2} \left(\frac{i_t}{i_{t-1}} - 1 \right)^2 \right] - 1 \right\} \right]$$

Optimal behavior of the capital goods producer gives

$$\lambda_t = \lambda_t \tilde{q}_t \left[1 - \frac{\kappa}{2} \left(\frac{i_t}{i_{t-1}} - 1 \right)^2 - \kappa \left(\frac{i_t}{i_{t-1}} - 1 \right) \frac{i_t}{i_{t-1}} \right] + \beta \mathbb{E}_t \left[\lambda_{t+1} \tilde{q}_{t+1} \kappa \left(\frac{i_{t+1}}{i_t} - 1 \right) \left(\frac{i_{t+1}}{i_t} \right)^2 \right] \tag{D.9}$$

and total dividend payments are

$$d_t^Q = i_t \left\{ \tilde{q}_t \left[1 - \frac{\kappa}{2} \left(\frac{i_t}{i_{t-1}} - 1 \right)^2 \right] - 1 \right\} \tag{D.10}$$

The retail block of the economy is fully summarized by equations (D.9) - (D.10).

Government. The government consumes resources g_t, unvalued by the representative household, subsidizes investment, and finances its expenditure on consumption and investment subsidies using lump-sum taxation:

$$g_t = \bar{g} \tag{D.11}$$

$$q_t = \tilde{q}_t(1 - \tau_t^i) \tag{D.12}$$

$$\tau_t = g_t + \int_{\mathcal{S}} \tau_t^i \tilde{q}_t i(s) d\mu(s) \tag{D.13}$$

Ricardian equivalence holds, so the precise path of financing of the given stream of government expenditure and investment subsidies is irrelevant. The nominal rate is set in accordance with a conventional Taylor rule:

$$\frac{1 + r_t^n}{1 + \bar{r}^n} = \left(\frac{1 + r_{t-1}^n}{1 + \bar{r}^n} \right)^{\rho_{tr}} \left[\left(\frac{1 + \pi_t}{1 + \bar{\pi}} \right)^{\phi_\pi} \left(\frac{y_t}{\bar{y}} \right)^{\phi_y} e^{u_t^m} \right]^{1 - \rho_{tr}} \tag{D.14}$$

where bars denote steady-state values. The government block of the economy is fully summarized by equations (D.11) - (D.14).

Exogenous Processes. Aggregate technology z_t, government spending g_t, the monetary policy shock u_t^m and government investment subsidies τ^i follow exogenously given paths. For the experiments in Section 4.2 we additionally generalize the firm leverage constraint Γ to depend on an exogenously evolving index of the tightness of financial constraints θ; details will be provided in Appendix D.3.

Market-Clearing. Output market-clearing requires that[19]

$$c_t + i_t + g_t = y_t \tag{D.15}$$

The debt market clears if

$$b_t = \int_{\mathcal{S}} b d\mu(s) \tag{D.16}$$

Finally labor market-clearing requires

$$\ell_t = \int_{\mathcal{S}} \ell(s) d\mu(s) \tag{D.17}$$

Equilibrium. We can now define an equilibrium in our full benchmark model.

Definition D.1. *Given an initial distribution μ_0 of firms over the state space $\mathcal{S}$, an interest rate on outstanding debt r_{-1} and a path of exogenous aggregates and government policies $\varepsilon = \{z, g, u^m, \tau^i, \theta\}$, a recursive competitive equilibrium is path of aggregate quantities $\{c_t, \ell_t, b_t, d_t, d_t^R, d_t^Q, \tau_t, \lambda_t, mc_t, y_t, i_t\}_{t=0}^{\infty}$ and prices $\{p_t^I, q_t, \tilde{q}_t, w_t, r_t, r_t^n, \pi_t, \pi_t^w\}_{t=0}^{\infty}$, value functions $\{v_t, v_t^l, v_t^d\}_{t=0}^{\infty}$, policy functions $\{\ell_t, i_t, k_t', b_t', d_t\}_{t=0}^{\infty}$ and distributions $\{\mu_t\}_{t=0}^{\infty}$ with the following properties:*

1. *Equations (D.1) - (D.17) hold.*

2. *Given the path of prices $\{p_t^I, w_t, r_t, q_t\}_{t=0}^{\infty}$ and aggregates $\{z_t, \theta_t\}_{t=0}^{\infty}$, the value functions $\{v_t^l, v_t^d\}_{t=0}^{\infty}$ and policy functions $\{\ell_t, i_t, k_t', b_t', d_t\}_{t=0}^{\infty}$ solve (A.1) - (A.3).*

3. *The distribution $\{\mu\}_{t=0}^{\infty}$ satisfies (A.4) with initial condition μ_0.*

4. *Aggregates $\{y_t, i_t, d_t\}_{t=0}^{\infty}$ are obtained from direct integration over the firm state space, using firm policy functions.*

Calibration. For our illustrations in Figure 3 and Figure 4 we need to choose particular parameter values for the non-production block of the economy. Since, for the parameterization of the production block, we want to replicate the original regressions in Zwick & Mahon (2017) as closely as possible, we choose a model period to correspond to a year. Given this choice, we choose all parameter values for the rest of the economy to be as consistent as possible with standard business-cycle models. Table 2 shows our parameter choices.

[19] For notational simplicity we assume that adjustment costs are rebated lump-sum back to households. Of course this assumption is inconsequential for our aggregation results.

Demand Block Parameterization

Parameter	Description	Value
β	Discount rate	1/1.04
h	Habit formation	0
γ	CRRA coefficient	1
φ	Frisch elasticity	∞
ϵ_p	Goods substitutability	10
θ_p	Price adjustment cost	40
ϵ_w	Wage substitutability	10
θ_w	Wage adjustment costs	100
κ	Aggregate K adjustment costs	0
ρ_{tr}	Taylor rule persistence	0.75
ϕ_π	Taylor rule inflation coefficient	1.5
ϕ_y	Taylor rule output coefficient	0

Table 2: Parameterization for rest of economy. Parameters are chosen to be largely in line with the standard New Keynesian literature.

Most parameter values are consistent with recent quantitative work in representative-agent business cycle models (Christiano et al., 2005; Smets & Wouters, 2007; Justiniano et al., 2010), with a couple of notable exceptions. First, we set the habit formation parameter h to 0. This choice is supposed to emphasize, in line with the results in Section 2 and somewhat contrary to Winberry (2018), that aggregation is a property of the production block itself, and not intrinsically related to any cyclical properties of prices. Second, we set the aggregate capital adjustment cost parameter κ to 0, consistent with empirical evidence on the high elasticity of capital goods supply (House & Shapiro, 2008; House et al., 2017). Third, we assume linear labor disutility (infinite Frisch elasticity) and a zero weight on output in the monetary authority's reaction function. These assumptions are made for simplicity and in line with Assumption 1, but are not central to our results.

To complete the model specification it remains to set various steady-state quantities and specify the exogenous processes. We normalize $\bar{w} = 1$, and then back out labor disutility χ residually to clear the labor market. We furthermore set the steady-state fraction of government consumption $\bar{g}/\bar{y}$ to 0.25. Finally, we assume that aggregate technology, government spending and monetary policy shocks follow AR(1) processes, with persistence $\rho_z = 0.75$,

$\rho_g = 0.45$ and $\rho_m = 0$, respectively. Our results are insensitive to these choices. We discuss the shock process for the tightness of financial constraints θ in Appendix D.3.

D.2 Monetary policy

FIRM BLOCK. We restrict the general production block (1) as follows. First, we assume a Cobb-Douglas production function, $f(k, \ell) = (k^\alpha \ell^{1-\alpha})^\nu$. Second, there are no financial frictions, so $\underline{d} = -\infty$ and $\Gamma(\bullet) = \infty$. Third, we consider a particular capital adjustment cost function $\phi(\bullet)$ of the form $\phi(k', k) = \tilde{\phi}(k, k' - (1-\delta)k)$ where

$$\tilde{\phi}(k, i) = \frac{\kappa}{2} \left(\frac{i - \delta k}{k} \right)^2 k - \vartheta i \mathbf{1}_{i<0} + \xi \mathbf{1}_{i \notin [-ak, ak]}$$

and where $\xi \overset{iid}{\sim} \text{uniform}[0, \bar{\xi}]$. This specification of adjustment costs is rich enough to nest most previous work; notably, it is as general as Cooper & Haltiwanger (2006), who argue that, to match certain micro investment data, it is necessary to allow for fixed costs, investment irreversibility, and quadratic adjustment costs, exactly as done here.

CALIBRATION. We split the set of model parameters into two blocks. The first block governs firm production. It contains the parameters determining the evolution of firm productivity – the support $\mathcal{E}$, the transition measure π, and the birth measure π_0 –, as well as those for the production function – the capital exponent α and the returns to scale parameter ν. The second block contains all parameters that govern real frictions to resource allocation across firms. Those include the capital depreciation rate δ, quadratic adjustment cost κ, the irreversibility coefficient ϑ, the fixed cost parameters $(\bar{\xi}, a)$, and initial capital k_0.

We set a model period to correspond to one year. For simplicity, a subset of the model parameters is fixed. Firm death π^d is set so that 6.5 per cent of firms enter and exit in any given period, bringing the ergodic age distribution in the model as close as possible to its empirical counterpart. The depreciation rate δ is set to imply an average investment-to-capital ratio of 6.5 per cent, a standard value in the literature. The coefficient on capital in the production function, α, as well as the returns to scale parameter, ν, are similarly set to conventional values (Khan & Thomas, 2013; Ottonello & Winberry, 2018). Given that neither conventional moments on firm investment nor our estimates of the investment sensitivity β are likely to be very informative about the firm-level productivity process (Clementi & Palazzo, 2016), we take log productivity to follow an AR(1) process, with parameter values coming from direct estimation on production function residuals.[20] This gives parameters (ρ, σ) for continuing firms; the productivity of entering firms is then drawn not from the implied stationary distribution, but instead from a shifted distribution $\log \mathcal{N}(\mu_0, \sigma_0)$.[21]

The remaining model parameters are disciplined as follows. First, following exactly the steps outlined in Appendix B.4, we replicate the experiment of Zwick & Mahon (2017) in our structural model, and target the resulting price elasticity estimand. Second, we add several conventional targets on firm investment behavior: the average investment rate $\mathbb{E}_s(i/k)$, the standard deviation of investment rates $\sigma_s(i/k)$, the investment spike rate $\mathbb{E}_s\left(\mathbf{1}_{i/k>0.2}\right)$, and the inaction rate $\mathbb{E}_s\left(\mathbf{1}_{|i|<0.01}\right)$. The full set of parameter values is summarized in Table 3.

Parameter Values

Parameter	**Description**	**Value**
Fixed Parameters		
$1-\xi$	Firm exit rate	0.065
δ	Depreciation rate	0.067
α	Capital share	0.310
ν	Returns to Scale	0.870
ρ	Productivity persistence	0.890
σ	Productivity dispersion	0.250
μ_0	Mean initial productivity	-0.375
σ_0	Initial productivity dispersion	0.330
Fitted Parameters		
κ	Quadratic adjustment costs	0.762
ϑ	Investment irreversibility	0.781
$\bar{\xi}$	Upper bound on fixed costs	0.450
a	Size of region without fixed costs	0.030
k_0	Capital of entrants	0.600

Table 3: Fitted parameters are chosen to match the calibration targets in Table 4.

[20] Direct estimates are provided, for example, in Cooper & Haltiwanger (2006), Foster et al. (2017), and David et al. (2018). Our process roughly falls in the middle of these estimates.

[21] We normalize mean productivity for continuing firms to 0. Mean productivity for entering firms is shifted relative to this normalization.

The most noteworthy feature of the overall calibration is the relatively large value of κ. These strong quadratic adjustment costs depart substantially from the calibration in Khan & Thomas (2008), and play a key role in dampening firm price sensitivity. In particular, a high value of κ is essential in generating a dampened price response coefficient and so crucially shapes the implied partial equilibrium response and general equilibrium adjustment.

Table 4 shows that the selected parameter values give a reasonable fit to our targets.

TARGETED MOMENTS

Target	**Data**	**Model**
Price Sensitivity		
Bonus depreciation estimand	2.890	2.984
Micro Investment		
Average investment rate	0.104	0.087
Std. of investment rates	0.160	0.147
Spike rate	0.144	0.108
Inaction rate	0.237	0.184
Employment Distribution		
Employment share of age-1 firms	0.016	0.028

Table 4: Price sensitivity following the discussion in Appendix B.4. Micro investment moments from annual firm-level IRS data, 1998 - 2010, as reported in Zwick & Mahon (2017), Appendix Table B.1. Employment shares of entrants from the Kaufmann survey and Cooper & Haltiwanger (2006).

D.3 Fiscal policy

FIRM BLOCK. We now restrict the general firm problem as follows. First, as for the first illustration, we assume a simple Cobb-Douglas production function. Second, and differently from before, we now allow for financial frictions, assuming that equity issuance is constrained by the lower bound on dividends $\underline{d}$, and that debt issuance must satisfy the earnings-based borrowing constraint

$$b_{jt} \leq \theta_t \times \pi_{jt}$$

Third, we assume the same real adjustment cost function as in Appendix D.2.

Calibration. We now split the model parameters into three blocks. The first two blocks are as before, while the new third block consists of parameters disciplining the severity of financial constraints; those include the equity issuance constraint $\underline{d}$, the borrowing constraint parameter θ and the initial debt endowments of newborn firms b_0. We again fix some parameters and estimate others; all fixed parameters from Table 3 are also used here, so we do not repeat them. We furthermore fix $\underline{d} = 0$, so firms can simply never issue equity. For the fitted parameters, we as before target evidence on the price sensitivity of investment as well as standard moments of the firm investment distribution. To discipline the severity of financial frictions, we simply match the direct evidence on corporate borrowing constraints in Lian & Ma (2018). These authors show that effective corporate borrowing constraints are well-characterized as earnings-based borrowing constraints, and provide direct evidence on the earnings scaling parameter θ. We complement this direct evidence with further restrictions on the initial size and financial position on newborn firms – firms that presumably face the most binding financial constraints. The full set of all new parameters is displayed in Table 5. Note that, for computational simplicity, we here do not allow for random fixed costs.

Parameter Values

Parameter	**Description**	**Value**
Fixed Parameters		
$\underline{d}$	Dividend constraint	0
Fitted Parameters		
κ	Quadratic adjustment costs	1.280
ϑ	Investment irreversibility	0.790
$\bar{\xi}$	Upper bound on fixed costs	0.00
a	Size of region without fixed costs	0.00
θ	Earnings-based borrowing constraint	3.000
k_0	Capital of entrants	0.420
b_0	Debt of entrants	0.180

Table 5: Fitted parameters are chosen to match the calibration targets in Table 6.

As before, the selected parameter values give a reasonable fit to our targets, even with the added restriction of no firm-level fixed adjustment costs.

Targeted Moments

Target	**Data**	**Model**
Price Sensitivity		
Bonus depreciation response	2.890	3.348
Micro Investment		
Average investment rate	0.104	0.136
Std. of investment rates	0.160	0.131
Spike rate	0.144	0.257
Inaction rate	0.237	0.205
Financial Frictions		
Earnings-based borrowing constraint	3.000	3.000
Entrants debt/output	1.280	1.501
Employment Distribution		
Employment share of age-1 firms	0.016	0.018

Table 6: Price sensitivity following the discussion in Appendix B.4. Micro investment moments from annual firm-level IRS data, 1998 - 2010, as reported in Zwick & Mahon (2017), Appendix Table B.1. Financial frictions directly following the estimates in Lian & Ma (2018). Debt and employment shares of entrants from the Kaufmann survey and Cooper & Haltiwanger (2006).

E Computational appendix

This appendix provides details on our computational routines. In Appendix E.1 we discuss computation of steady-state equilibria, and in Appendix E.2 we turn to computation of perfect-foresight transition dynamics.

E.1 Steady state

Value and Policy Functions. We solve the firm problem (A.1) -(A.3) using collocation methods. The idiosyncratic productivity process is discretized using the standard Rouwenhorst method. For debt and capital, we set up a grid of collocation nodes on the state space $\mathcal{S}$. For capital, we use a cubic spline on a grid $[k_{min}, k_{max}]$, where k_{min} and k_{max} are, respectively, the minimal and maximal capital levels with strictly positive density in a variant of our model without financial frictions.[22] For debt, we use a linear spline on $[b_{min}, b_{max}]$, where b_{min} is computed from the minimal savings policy (Khan & Thomas, 2013) and b_{max} is set large enough so that no firms are at the upper state boundary.

Given a guess for the spline coefficients, we then iterate towards a set of coefficients that solve the firm's Bellman equation.[23] In a given iteration, we use golden search to find the optimal capital policy, given future debt set residually to ensure zero dividends. If, given the state triplet (e, k, b), the firm is already perpetually unconstrained, we instead set dividends to correspond to the minimal savings policy, and set capital according to the unconstrained optimum. Given optimal policies, we find new spline coefficients. Rather than re-optimizing at every step, we use intermediate Howard improvement steps. This procedure is iterated until convergence.

Stationary Distribution. We construct the stationary distribution through iteration on firm policy functions over the state space, as in Ahn et al. (2017). For accuracy, we compute firm decisions on a finer grid for capital and debt, given our previous approximation of continuation values on the original coarse grid. Using the spline bases, we map the derived policy functions – which generically do not map onto the fine grids – into weighted averages on the different grid points, ensuring that aggregates computed from the stationary distribution

[22]For the model without financial frictions, we use as an initial guess of the boundaries the minimal and maximal capital levels in a model without any micro frictions. These bounds can be computed analytically.

[23]Infeasible regions of the state space are penalized. The penalty value is set high enough so that even higher values do not materially affect our results.

will be unbiased (Gavazza et al., 2018). We adjust the full transition rule to take into account firm death and exit. Given a final transition matrix $\mathbf{Q}$, we can easily compute the stationary distribution as a solution to

$$Q' \times \bar{\mu} = \bar{\mu}$$

where as usual $\mathbf{1}' \times \bar{\mu} = 1$.

E.2 Dynamics

Throughout, we truncate the computation of dynamic transition paths at large T, where T is chosen large enough to ensure transition back to steady state. In our benchmark model, we only need to find the equilibrium sequence $\tilde{y}_t$ to solve a generalized analogue of (7),

$$\mathbf{y}(\mathbf{p}(\tilde{\mathbf{y}});\boldsymbol{\varepsilon}) - \mathbf{i}(\mathbf{p}(\tilde{\mathbf{y}});\boldsymbol{\varepsilon}) = \tilde{\mathbf{y}}$$

For this we proceed using a quasi-Newton method. First, given $\tilde{y}_t = \bar{\tilde{y}}$ for all t, we solve the heterogeneous-firm block of the economy once to get the partial equilibrium effect $\tilde{\mathbf{y}}_\varepsilon$ *for a particular shock*. Second, using finite differences for $\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{p}}$ and automatic differentiation for $\frac{\partial \mathbf{p}}{\partial \tilde{\mathbf{y}}}$, we obtain an approximation of $\mathcal{G}$ *for all structural shocks*, evaluated at steady-state prices. We then obtain a new guess for $\tilde{\mathbf{y}}$ via a quasi-Newton step, using our approximation of $\mathcal{G}$ throughout:

$$\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_k - (I - \mathcal{G})^{-1} \times (\tilde{\mathbf{y}}_k - \tilde{\mathbf{y}}(\mathbf{p}(\tilde{\mathbf{y}}_k);\varepsilon)) \tag{E.1}$$

where $\tilde{\mathbf{y}}_0 = \bar{\tilde{\mathbf{y}}}$ and $\tilde{\mathbf{y}}(\mathbf{p}(\tilde{\mathbf{y}}_0);\varepsilon) = \tilde{\mathbf{y}}_\varepsilon$. We iterate this procedure until convergence.

Conditional Shocks. For our experiments on state dependence, we consider transition paths after two unexpected shocks. The original transition path after the first shock is solved as before. Then, with a delay of $\ell \geq 0$ periods, a second shock hits. Taking as our initial condition the current distribution μ_ℓ, we then solve for the effects of both shocks jointly. We then obtain the partialled-out effect of the second structural shock by subtracting the impulse responses to the first shock only from the joint transition path.

Extended Models. For the extended models discussed in Appendix A.1 the solution technique proceeds in almost exactly the same way. We now simply need to clear more markets, so the computation of the updating matrix $\mathcal{G}$ becomes more time-consuming.