Aggregate Dynamics in Lumpy Economies

Isaac Baley
Andres Blanco

May 2019

Aggregate Dynamics in Lumpy Economies[*]

Isaac Baley[†]

UPF, CREi, and Barcelona GSE

Andrés Blanco [‡]

University of Michigan

May 27, 2019

Abstract

In economies with lumpy microeconomic adjustment, we establish structural relationships between the dynamics of the cross-sectional distribution of agents and its steady-state counterpart and discipline these relationships using micro data. Applying our methodology to firm lumpy investment, we discover that the dynamics of aggregate capital are structurally linked to two cross-sectional moments of the capital-to-productivity ratio: its dispersion and its covariance with the time elapsed since the last adjustment. We compute these sufficient statistics using plant–level data on the size and frequency of investments. We find that, in order to explain investment dynamics, the benchmark model with fixed adjustment costs must also feature a precise combination of irreversibility and random opportunities of free adjustment.

JEL: D30, D80, E20, E30

Keywords: inaction, lumpiness, transitional dynamics, non convex adjustment costs, sufficient statistics, firm investment, adjustment hazard, Ss models.

---

1 Introduction

Lumpiness in microeconomic adjustment is pervasive in many economic environments. Capital investment, labor hiring and firing, inventory management, consumption of durable goods, price setting, portfolio choice, and many other economic decisions faced by firms and households are characterized by periods of inaction followed by bursts of activity. How does lumpiness in microeconomic adjustment affect aggregate dynamics? How persistent are the effects of an aggregate shock or a change in government policy in lumpy environments? Understanding these issues is key for the design and implementation of policies aimed at stabilizing the business cycle or generating growth.

This paper presents a new approach to study aggregate dynamics in lumpy economies. We consider environments with ex ante identical agents that make decisions subject to idiosyncratic shocks and adjustment frictions. These frictions may take the form of non convex adjustment costs, random or fixed dates of adjustment, among others. Aggregate dynamics in this type of environment require handling the cross-sectional distribution of agents, which is a highly dimensional and complicated object. We propose a 'sufficient statistics' approach for characterizing aggregate dynamics that consists of (i) representing the *dynamics* of the distribution as a function of a few *steady-state* cross-sectional moments, and (ii) measuring these steady-state moments through *micro data on adjustments*. Our sufficient statistics approach provides researchers with a unique set of moments to be targeted by lumpy models and guides the empirical efforts in providing the statistics that are most informative for the theory.

Our first result characterizes the dynamics of the distribution in the following way. Starting from an initial distribution close to the ergodic one, we characterize the convergence of any moment of such distribution toward its ergodic counterpart. Following Álvarez and Lippi (2014), we focus the analysis on the cumulative impulse–response function, or CIR, which equals the cumulative deviations in the moment of interest with respect to its steady-state value, i.e., the area under the IRF. The CIR is a useful metric: It summarizes the impact effect and persistence of distributional dynamics following a once and for all aggregate shock, it eases comparison across models, and it can be interpreted as a "multiplier" of aggregate shocks.[1] We show analytically that the CIR can be expressed as a linear combination of steady-state cross-sectional moments. While the exact mapping depends on the moment being tracked and some features of the environment (for example, time- vs. state- dependent adjustments), we characterize the CIR for widely used models of inaction in a variety of economic applications.

The logic of the previous result is that steady-state moments are informative about agents' responsiveness to idiosyncratic shocks, and therefore are good predictors of the effect of lumpiness on the speed of convergence of the distribution. As an example, consider a monetary economy with lumpy price-setting and no inflation. Then, substantial steady-state dispersion in firm-level markups relative to the volatility of idiosyncratic productivity shocks must necessarily reflect important adjustment frictions, which in turn generate unresponsive adjustments to aggregate nominal shocks and persistent real output effects. In this simple example, the ratio of steady-state markup dispersion to idiosyncratic volatility is a sufficient

---

[1] Álvarez, Le Bihan and Lippi (2016) and Baley and Blanco (2019) also use the CIR to compare the real effect of monetary policy shocks across price-setting models. The analysis of the CIR is closely related to the "marginal response function" by Borovička, Hansen and Scheinkman (2014).

statistic that characterizes the dynamics of average markups following a monetary shock.[2]

Our second result shows how to recover these sufficient statistics—the steady-state cross-sectional moments of (potentially unobserved) variables—using information on observable actions available in micro datasets. To fix ideas, assume an economy with lumpy capital adjustment and idiosyncratic productivity shocks and suppose there is interest in recovering a measure of misallocation, that is, the steady-state dispersion in the capital-to-productivity ratio. The idea consists of assuming a continuous process for idiosyncratic productivity that hit firms during periods of inaction and a minimal structure for the adjustment policy, and then using the information revealed through firms' observed investments to back out the level of misallocation. The minimal assumption requires that adjustment frictions deliver a constant reset point, that is, that all firms set the same capital-to-productivity ratio when they decide to invest.[3] In this way, our theory shows how one can recover capital misallocation using data on investment rates exclusively.

Application: Aggregate capital dynamics. To illustrate our framework's usefulness, we consider an application to aggregate capital dynamics. For this purpose, we set up a model of lumpy investment in the spirit of Caballero and Engel (1999) and the related literature.[4] Firms are subject to idiosyncratic productivity shocks, a common drift due to depreciation and productivity growth, and a general model of adjustment costs. Firms must pay a fixed cost to invest, which could be different for positive and negative investments and depend on firms' size. In addition, firms face random opportunities of free adjustment.

We demonstrate analytically that, in order to explain aggregate capital dynamics as measured via the CIR, such a model must match two steady-state moments related to capital misallocation and the distribution of times since the last adjustment. Concretely, defining the capital gap as the log of the capital-to-productivity ratio relative to the steady-state cross-sectional average, the CIR for the first moment of the capital gap is given by two steady-state moments: the *variance* of capital gaps—a measure of misallocation—and the *covariance* between capital gaps and capital gaps' age (the time elapsed since last adjustment). Moreover, we show numerically that these two sufficient statistics are also quantitative determinants in models with richer adjustment cost structures (e.g., full irreversibility, U-shaped hazards).

Given the mapping from the CIR to ergodic moments of capital gaps, we apply the second part of the theory to infer the variance and covariance terms from micro data on adjustments. Using plant-level data from Chile and Colombia, we measure these two key moments using the frequency and size of investments (notably, the covariance term had never been computed before) and discover two interesting regularities. First, the variance of capital gaps is significantly larger than the volatility of idiosyncratic productivity, a telltale sign that policies are insensitive to the capital gap and favoring random opportunities of free adjustment. Second, the covariance between capital gaps and capital gaps' age is positive, which means that firms that have not invested in a long time have large capital-to-productivity ratios, even in the presence of depreciation and productivity growth. This is evidence that adjustment frictions are asymmetric.

---

[2]Blanco (2019) shows quantitatively that the slope of the Phillips curve is linked to ergodic relative price dispersion.

[3]This feature is present in models with non-convex adjustment costs and random or fixed adjustment dates, among others.

[4]Similar environments have been studied by Dixit and Pindyck (1994), Bertola and Caballero (1994) and Caballero, Engel and Haltiwanger (1995), and more recently in quantitative general equilibrium by Veracierto (2002), Thomas (2002), Gourio and Kashyap (2007), Khan and Thomas (2008), Bachmann, Caballero and Engel (2013), and Winberry (2019).

Our theory shows how missing the previous two moments in the data leads to a wrong assessment of the role of micro-level lumpiness for aggregate dynamics. For example, a standard fixed-cost model dramatically misses these moments and thus can only generate up to 1% of the CIR inferred from the data; the widespread model with a U-shaped probability of adjustment slightly increases the fit up to 14% of the CIR. By adding frequent opportunities of free adjustment and strong downward irreversibility, the model can generate up to 75% of the CIR. We conclude that our methodology can aid researchers in understanding the deep links that exist between the nature of adjustment costs, capital misallocation, and the dynamics of aggregate investment.

Advantages and limitations. Our approach allows us to characterize the CIR for any moment of the cross-sectional distribution and other continuous functions of these moments. This is useful as many applications require tracking the dynamics of second, third, or higher moments of cross-sectional distributions. Moreover, we can accommodate transitions starting from various initial distributions consisting of small perturbations around the steady-state, such as horizontal shifts (first-moment shocks) or mean-preserving spreads (second-moment shocks). In the same vein, we can accommodate a variety of continuous stochastic processes for the state (e.g., with drift, without drift, mean-reverting). While the mappings between data and ergodic moments clearly depend on the stochastic process assumed, the theory imposes cross-equation restrictions that allow us to either validate or reject the assumptions.

One limitation of our framework is that we can only characterize the CIR but not the complete IRF. The reason is that we need to consider the complete transitional dynamics in order to exploit the ergodic properties of the environment (i.e., exchange the time series of the cross-sectional distribution with a cross-sectional distribution of individual stopping-time policies). Extending the theory to characterize the full IRF would allow us to better understand the role of lumpiness for aggregate dynamics.

A second limitation is that our analysis takes as a premise that the steady-state policies hold along the transition path. This assumption is valid as long as the general equilibrium feedback from the distribution to individual policies though prices is quantitatively insignificant. There are several general equilibrium frameworks in which this is the case, either theoretically or quantitatively.[5] But when general equilibrium effects are quantitatively relevant, our framework does not fully characterize aggregate dynamics. Nevertheless, the tools developed in this paper are still informative about the role of lumpiness in richer general equilibrium models and serve as a guide to study one important dimension of the economic environment. In order to minimize these concerns, our empirical application considers data from small open economies, Chile and Colombia, for which the interest rate is mainly determined outside their borders.

---

[5]In the sticky-price models of Woodford (2009), Golosov and Lucas (2007), and the vast literature that builds on them, equilibrium nominal wages respond one-to-one to money supply shocks and equilibrium real wages have a negligible effect on firms' policies. Similar models are also used to study real exchange rate dynamics by Carvalho and Nechio (2011) and Blanco and Cravino (2018). In investment, Bachmann, Caballero and Engel (2013) and Winberry (2019), building on Khan and Thomas (2008), show that partial equilibrium dynamics are not undone by general equilibrium effects whenever the model matches properties of the data. the properties of wages and interest rates in the data. For example, the general equilibrium feedback is small when matching the cyclical behavior of US interest rates (Winberry, 2019) or the conditional heteroscedasticity of the investment rates (Bachmann, Caballero and Engel, 2013). Online Appendix C describes some of these frameworks.

Contributions to the literature. We highlight three contributions with respect to previous work. Our first contribution is to provide sufficient statistics that capture the role of micro lumpiness for aggregate dynamics that are valid across a variety of models. Álvarez, Le Bihan and Lippi (2016) provide a first step in this direction. They show that in a large class of price-setting models, the CIR of average markup gaps—a measure of the real effects of a monetary shock—is proportional to the kurtosis of price changes times the average price duration.[6] In related work, Carvalho and Schwartzman (2015) and Álvarez, Lippi and Paciello (2016) show that in time-dependent models the relevant sufficient statistic is the average price duration times its coefficient of variation. These results are restricted to economies with zero inflation, symmetric policies, and the transitions of the first moment of the distribution. Our contribution lies in establishing steady-state cross-sectional moments as an intermediate link between the CIR and observable actions. With this strategy, we can extend the applicability of the sufficient statistics approach to richer economic environments featuring drift and asymmetric policies and study the transitions of higher order moments.

The sufficient statistics approach considers a flexible formulation that permits us to remain agnostic about the nature of adjustment costs, giving the data the biggest role in informing about it. For instance, our work speaks to the debate about the size of the adjustment frictions in the quantitative investment literature.[7] By matching different conditional and unconditional moments generated by a particular model to their empirical counterparts, the literature has obtained values for fixed costs that can differ up to various orders of magnitude and reach opposite conclusions. We establish the two moments that these models need to match to explain investment dynamics and compute the first empirical estimates. In the same spirit, Elsby, Ratner and Michaels (2019) assess the role of labor market frictions for aggregate dynamics in terms of a sufficient statistic for labor flows across firms, which can be measured using establishment panel data on employment.

Our second contribution is to strengthen the bridge between two literatures that study lumpy economies with different objectives and methodologies. The first aims at understanding the role of lumpiness for aggregate dynamics and has a long tradition, see Caplin and Spulber (1987), Caplin and Leahy (1991, 1997), and Caballero and Engel (1991) for early work. The second literature aims at quantifying welfare losses and inefficiencies in steady-state, for example, Álvarez, Beraja, Gonzalez-Rozada and Neumeyer (2018) in the context of price-setting and Asker, Collard-Wexler and De Loecker (2014) in investment. To our best knowledge, this paper is the first to show theoretically that structural links exist between aggregate dynamics and the steady-state cross-sectional moments of lumpy economies. We believe our tools can engage researchers to establish deeper links across these fields and exploit the connections between these two dimensions of the same environment.

Lastly, our third contribution is to provide an alternative methodology for recovering *gaps* in lumpy environments, which are understood as the difference between an agent's choice under frictions relative to an optimal frictionless or static target. Since Hamermesh (1989), there have been several approaches to recover these gaps. In labor, Caballero, Engel and Haltiwanger (1997) use hours worked to recover the

---

[6]This sufficient statistic is valid for the price-setting models in Taylor (1980), Calvo (1983), Reis (2006), Golosov and Lucas (2007), Nakamura and Steinsson (2010), Midrigan (2011), Álvarez and Lippi (2014), among other.

[7]See Thomas (2002) and its reassessment by Gourio and Kashyap (2007), as well as Khan and Thomas (2008) and its reassessment by Bachmann, Caballero and Engel (2013) and Winberry (2019). Footnote 4 gives more references.

labor gap (see Cooper and Willis (2004) and Bayer (2009) for a discussion). In pricing, Campbell and Eden (2014) use the average price across the same retailer in the same city to recover the markup gap. In development, Hsieh and Klenow (2009) use value added over capital to recover the capital gap. Our methodology proposes an alternative way to directly compute cross-sectional moments of the gaps and to estimate the parameters of the assumed stochastic process using observable micro data on adjustments.

Structure. Section 2 presents a standard model of lumpy investment that allows us to introduce the objects of interest. Section 3 develops the theory. Section 4 applies the theory using micro-level data, and Section 5 generalizes and extends the results.

2 A Model of Lumpy Investment

This section describes the economic environment, built in the spirit of Caballero and Engel (1999), within which we develop and apply the theory.

2.1 Environment

Time is continuous and infinite. A representative household and a continuum of ex ante identical firms live in the economy. There is no aggregate uncertainty, and firms face idiosyncratic shocks to productivity and capital adjustment costs. We denote with $\omega \in \Omega$ the full history of shocks and consider $(\Omega, P, \mathcal{F})$ to be a probability space equipped with the filtration $\mathcal{F} = \{\mathcal{F}_t : t \geq 0\}$. We use the notation $g_{\omega,t} : \Omega \times \mathbb{R} \to \mathbb{R}$ to denote an adapted process (a function $\mathcal{F}_t$-measurable for any $t \geq 0$) and $\mathbb{E}[g_{\omega,t}]$ to denote its expectation under P.

Household. The household chooses the stochastic process for consumption to maximize its expected utility subject to a budget constraint. The household's problem is given by

$$\int_0^\infty e^{-\rho t} C_t dt, \quad \text{subject to} \quad \int_0^\infty Q_t \left(C_t - \Pi_t \right) dt = 0, \tag{1}$$

where C_t denotes the household's consumption, $\Pi_t \equiv \mathbb{E}[\pi_{\omega,t}]$ denotes firms' aggregate profits, and Q_t is the Arrow–Debreu time-zero price. Due to the linearity of preferences, prices are equal to $Q_t = e^{-\rho t}$ and are independent of the aggregate state (the distribution of firm idiosyncratic states, described next).

Firms. Firms operate in competitive markets. They produce output of a uniform good using capital (k), subject to idiosyncratic productivity (e). Capital between adjustments depreciates at a constant rate $\zeta > 0$. The production function features decreasing returns to scale $\alpha < 1$. For every capital adjustment, a firm pays an adjustment cost proportional to its size and it may be subject to partial irreversibility. The adjustment cost has the general form

$$\tilde{\theta}_{\omega,t} \equiv \theta_{\omega,t} \mathcal{J} \left(e_{\omega,t}, k_{\omega,t^-}, \mathbb{1}_{\{\Delta k_{\omega,t} > 0\}} \right), \tag{2}$$

where $\mathcal{J}(\cdot)$ is homogeneous of degree one in the first two arguments (k_{ω,t^-} denotes the left limit of the capital stock) and may depend on the sign of the investment ($\mathbb{1}_{\{\Delta k_{\omega,t}>0\}}$ takes the value 1 if investment is positive and 0 otherwise). The component $\theta_{\omega,t}$ follows

$$\theta_{\omega,t} = \begin{cases} \xi & \text{if } dN_t = 1 \\ \theta & \text{otherwise} \end{cases}, \tag{3}$$

where N_t is a Poisson counter with arrival rate λ and $\xi \sim H_\xi$ with support $[0,\theta]$.[8]

A firm chooses a sequence of adjustment dates $(T_{\omega,i})$ and investment $(\Delta k_{\omega,T_{\omega,i}} = k_{\omega,T_{\omega,i}} - k_{\omega,T_{\omega,i}^-})$ to maximize its expected stream of profits, discounted at the household's discount factor Q_t:

$$\max_{\{T_{\omega,i}, \Delta k_{T_{\omega,i}}\}_{i=1}^{\infty}} \mathbb{E}\left[\int_0^\infty Q_s e_{\omega,s}^{1-\alpha} k_{\omega,s}^\alpha \, ds - \sum_{i=1}^{\infty} Q_{T_{\omega,i}}\left(\tilde{\theta}_{\omega,T_{\omega,i}} + \Delta k_{\omega,T_{\omega,i}}\right)\right], \tag{4}$$

where productivity and capital follow:

$$\log(e_{\omega,t}) = \log(e_{\omega,0}) + \mu t + \sigma W_{\omega,t}, \tag{5}$$

$$\log(k_{\omega,t}) = \log(k_{\omega,0}) - \zeta t + \sum_{i:T_{\omega,i}\leq t} \log(1 + i_{\omega,T_{\omega,i}}). \tag{6}$$

Here $\mu > 0$ is a common trend, idiosyncratic productivity shocks $W_{\omega,t}$ follow a Wiener process with volatility $\sigma > 0$, and $i_{\omega,T_{\omega,i}} = \Delta k_{\omega,T_{\omega,i}}/k_{T_{\omega,i}^-}$ denotes the investment rate at the i-th adjustment.

Aggregate feasibility. Aggregate output Y_t is used for the household's consumption C_t and aggregate investment I_t, which includes capital adjustment costs:

$$\underbrace{\mathbb{E}\left[e_{\omega,t}^{1-\alpha} k_{\omega,t}^\alpha\right]}_{Y_t} = C_t + \underbrace{\mathbb{E}\left[\mathbb{1}_{\{\tau,t\}}\left(\tilde{\theta}_{\omega,t} + \Delta k_{\omega,t}\right)\right]}_{I_t}. \tag{7}$$

Here $\mathbb{1}_{\{\tau,t\}} = \{\omega : \exists i \text{ such that } T_{\omega,i} = t\}$ indicates the set of adjusters.

Equilibrium. Given an initial distribution of $\{k_{\omega,0}, e_{\omega,0}\}$, an equilibrium is a set of stochastic processes for Arrow-Debreu prices $\{Q_t\}$, the household's consumption policy $\{C_t\}$, and the firms' investment policies $\{T_{\omega,i}, \Delta k_{\omega,T_{\omega,i}}\}$ such that:

(i) Given prices $\{Q_t\}$, $\{C_t\}$ solves the household's problem (1).

(ii) Given $\{Q_t\}$ and the stochastic processes for productivity and capital adjustment costs, $\{T_{\omega,i}, \Delta k_{\omega,T_{\omega,i}}\}$ solve the firm's investment problem (4).

(iii) The goods market clears in (7).

---

[8]We omit here a more general formulation that encompasses richer time-dependent models to simplify exposition of the model. See below for the formulation of $\theta_{\omega,t}$ to generate time-dependent models and Section 5 for the extension to more general state spaces and policy functions.

2.2 Redefining the state: Capital gap

Capital gap. Since the economy features growth, we first redefine the firms' state to ensure stationarity. To this end, we define three variables. First, we define a firm's log capital–productivity ratio $\hat{k}_{\omega,t} \equiv \log(k_{\omega,t}/e_{\omega,t})$. Second, we define the steady-state average capital–productivity ratio $\hat{k}_{ss} \equiv \mathbb{E}\left[\log(k_\omega/e_\omega)\right]$. The notation without time index t refers to moments computed with the steady-state distribution. Lastly, we define the *capital gap* as the firm's capital–productivity ratio minus its steady-state average

$$x_{\omega,t} \equiv \hat{k}_{\omega,t} - \hat{k}_{ss}. \tag{8}$$

While the distribution of productivity may diverge, the distribution of capital gaps is stationary and has mean zero because of the normalization. Importantly, the construction of capital gap does not use the static optimal choice as it is commonly done in the literature, but the economy's average log capital–productivity ratio.

Law of motion of capital gaps. *Uncontrolled* capital gaps—not considering any investments—follow the process

$$d\tilde{x}_{\omega,t} = \nu dt + \sigma dW_{\omega,t}, \tag{9}$$

where we use tildes to show explicitly that these variables evolve exogenously. For convenience, we define the total drift of capital gaps as $\nu \equiv -(\zeta + \mu) < 0$, which is negative and includes depreciation and the common productivity trend. Initial conditions $\tilde{x}_{\omega,0}$ are exogenously given. By the discussion above, the initial condition of the uncontrolled capital gap is $\tilde{x}_{\omega,0} = \hat{k}_{\omega,0} - \hat{k}_{ss}$.

In contrast, *controlled* capital gaps—which take into account the adjustment policy— evolve as

$$x_{\omega,t} = \tilde{x}_{\omega,t} + \sum_{T_{\omega,i} \leq t} \Delta x_{\omega,T_{\omega,i}}, \tag{10}$$

where the adjustment dates $(T_{\omega,i})$ and the changes in capital gaps $(\Delta x_{\omega,T_{\omega,i}})$ solve the firm's problem in (4). Due to the continuity assumption for the idiosyncratic productivity, the *changes in the capital gap* equal the changes in the *capital stock* or the investment rate $i_{\omega,T_{\omega,i}} \equiv \Delta k_{\omega,T_{\omega,i}}/k_{\omega,T_{\omega,i}^-}$ as follows:

$$\Delta x_{\omega,T_{\omega,i}} = \lim_{t\uparrow T_{\omega,i}} \log(k_{\omega,T_{\omega,i}}/k_{\omega,t}) - \underbrace{\lim_{t\uparrow T_{\omega,i}} \log(e_{\omega,T_{\omega,i}}/e_{\omega,t})}_{=0} = \log(k_{\omega,T_{\omega,i}}/k_{\omega,T_{\omega,i}^-}) = \log\left(1 + i_{\omega,T_{\omega,i}}\right). \tag{11}$$

Reset capital gap. Homotheticity and *iid* behavior in the adjustment cost imply that the *reset capital gap* $\hat{x}$—the new capital gap conditional on adjustment—is constant across agents and time.[9] Thus, the reset gap is memoryless, in the sense that once an agent adjusts, it incorporates all of the accumulated idiosyncratic shocks into its investment:

$$\hat{x} = x_{\omega,T_{\omega,i}}. \tag{12}$$

---

[9]Homogeneous resets are also the result of firms' being ex ante identical, an assumption that is relaxed in Section 5.2.

With the normalization explained above, $\hat{x}$ is the capital–productivity ratio of adjusting firms *relative* to the average ratio in the economy, and *not* relative to a static optimal value.

Duration of inaction and capital gap age. Finally, we define two related but different concepts. First, the duration of an inaction period—denoted by τ—is the difference between two consecutive adjustment dates:

$$\tau_{\omega,i} \equiv T_{\omega,i+1} - T_{\omega,i}. \tag{13}$$

Second, the *age* of the capital gap—denoted by a—is the time elapsed since its last adjustment:

$$a_{\omega,t} = t - \max\{T_{\omega,i} : T_{\omega,i} \le t\}. \tag{14}$$

After each adjustment, *age* is reset to zero. The joint stochastic process of (x, a) does matter for aggregate dynamics, and therefore we also include a as a state.

2.3 From adjustment cost structures to policies

Our specification for capital adjustment costs $\tilde{\theta}_{\omega,t}$ stated in (2) spans a wide set of models. The advantage of considering a flexible formulation is that it permits us to remain agnostic about the true nature of adjustment costs and allows the data to inform us about it.

The first class refers to ***fully state-dependent models***, in which firms always face the same *constant fixed cost* $\xi = \theta$, i.e., $H_\xi(\theta) = 1$ and $H_\xi(\xi) = 0$ for all $\xi < \theta$. This structure is considered by Caballero and Engel (1991); it features an inaction region $[\underline{x}, \overline{x}]$ and produces the following sequence of duration of inaction:

$$\tau_{\omega,i} = \inf_{t \ge 0} \left\{ t : x_{\omega,T_{\omega,i}+t} \notin [\underline{x}, \overline{x}] \right\}. \tag{15}$$

The second class refers to ***random fixed-cost models***, which are hybrid models with both time- and state-dependence in the adjustment cost structure. Consider H_ξ to be non degenerate, so that firms draw new values of the fixed cost. Let $\tilde{N}_{\omega,t}$ be a Poisson counter with a general hazard rate $\Lambda(x_{\omega,t^-})$. Then adjustments dates are given by

$$\tau_{\omega,i} = \inf_{t \ge 0} \left\{ t : x_{\omega,T_{\omega,i}+t} \notin [\underline{x}, \overline{x}] \ \text{ or } \ d\tilde{N}_{\omega,T_{\omega,i}+t} \ge 1 \right\}. \tag{16}$$

A broadly used example, proposed by Caballero and Engel (1999), takes $H_\xi \sim Uniform[0,\theta]$; it implies a U–shaped adjustment hazard that is equal to zero at the reset capital-productivity ratio and increasing in the distance. Our theory is more general, as it can accommodate non monotonic hazard rates as well as a non zero hazard at the reset point. Within this class, if H_ξ is degenerate at zero, then firms face either a positive or a zero adjustment cost. We label this specification as the ***Bernoulli fixed cost***; it implies a constant adjustment hazard inside the inaction region and is the workhorse model of the paper, as it allows for a sharp characterization.[10] The sequence of duration for the Bernoulli fixed-cost model follows (16) with $\Lambda(x_{\omega,t^-}) = \lambda$.

---

[10]In the price-setting literature, this is labeled the CalvoPlus model by Nakamura and Steinsson (2010).

Lastly, we consider **fully time-dependent models**. This case is interesting for two reasons: Many applications (e.g., pricing and portfolio choice) consider time-dependent costs; moreover, it allows us to better explain economic intuitions. These models can be summarized by a sequence of duration of inaction given by

$$\tau_{\omega,i} \ = \ u_{\omega,i}, \qquad iid \sim cdf \ H_u(u). \tag{17}$$

If, in addition, u is a constant—i.e., $H_u(u) \ = \ 0$ for all $u < T$ and $H_u(T) \ = \ 1$—then firms adjust at the fixed date T. If H_u follows an exponential distribution with parameter λ, then firms adjust at random iid dates.[11]

Asymmetric costs and irreversibilities. The third argument in $\mathcal{J}(e_{\omega,t}, k_{\omega,t^-}, \mathbb{1}_{\{\Delta k_{\omega,t}>0\}})$ allows us to consider asymmetries in adjustments costs. If $\mathcal{J}(\cdot,\cdot,1) = \mathcal{J}(\cdot,\cdot,0)$, then any policy asymmetry (inaction region or adjustment hazard) must be a direct consequence of asymmetry in the static payoff or drift in the capital–productivity ratio. This property breaks whenever $\mathcal{J}(\cdot,\cdot,1) \neq \mathcal{J}(\cdot,\cdot,0)$, as asymmetric policies also arise by setting different costs for positive and negative investments, for example, by setting $\mathcal{J}(\cdot) = e\theta(1 + \theta^-(1 - \mathbb{1}_{\{\Delta k_{\omega,t}>0\}}))$; the extreme case $\theta^- = \infty$ generates full downward irreversibility, as in Veracierto (2002).

In the next sections, we focus the analysis on the families specified above.[12]

2.4 Steady-state and aggregate dynamics

We are interested in characterizing the deviations of aggregate variables from the steady-state. First, we show that aggregate variables can be expressed as moments of the capital gap distribution. Then, we define steady-state moments and our notion of aggregate dynamics through the cumulative impulse response, or CIR.

Aggregate variables. Here we show a few examples of how deviations of aggregate variables can be expressed in terms of capital gaps. Take aggregate capital detrended by productivity, denoted with $\hat{K}_t \equiv \mathbb{E}\left[\log(k_{\omega,t})\right] - \mathbb{E}\left[\log(e_{\omega,t})\right] = \mathbb{E}\left[\log(k_{\omega,t}/e_{\omega,t})\right]$. Then, the aggregate capital's deviation from steady-state is equal to the average capital gap $\mathbb{E}[x_{\omega,t}]$:

$$\hat{K}_t - \hat{K}_{ss} \ = \ \mathbb{E}\left[\log(k_{\omega,t}/e_{\omega,t})\right] - \mathbb{E}\left[\log(k_\omega/e_\omega)\right] \ = \ \mathbb{E}\left[\hat{k}_{\omega,t}\right] - \hat{k}_{ss} \ = \ \mathbb{E}[x_t]. \tag{18}$$

Notice that in this normalization, we first centralize the capital-gap distribution around its steady-state average and then aggregate across firms. Output deviations from steady-state can also be expressed in

---

[11]The micro-foundation for these models is $\theta_{\omega,t} = 0$ if $\sum_{j\,=\,1}^{i} u_{\omega,i} = t$ for some i or $\theta_{\omega,t} = \infty$ otherwise.

[12]Let us compare our environment with the lumpy investment models cited before. First, we do not consider labor as a production factor. Given that we consider a partial equilibrium setting and the labor decision is usually static, this assumption is innocuous, as adding labor only affects the value of the output-capital elasticity. Second, all of the investments in our model, regardless of their size, are subject to adjustment costs (in the language of the literature, we do not consider *unconstrained* investments). Quantitatively, this assumption does not play a big role in the dynamics, as in the calibration most investments are constrained due to the large size of idiosyncratic shocks. Lastly, we consider a random-walk process for idiosyncratic productivity instead of a mean-reverting process. This assumption is done to simplify the exposition at this stage and is relaxed in Section 5.2. Online Appendix F explores the role of these alternative assumptions.

terms of moments of normalized capital gaps: $\hat{Y}_t = \alpha\mathbb{E}[x_t]$. Misallocation, defined as the variance of log marginal products of capital, can be expressed in terms of the variance of capital gaps: $\mathbb{V}ar[\log(mpk_t)] = (1-\alpha)^2\mathbb{V}ar[x_t]$. Overall, the previous analysis illustrates how to shift the focus from aggregate variables to moments of the capital gaps.

Steady-state moments. Consider the steady-state distribution of capital gaps and age, denoted by $F(x,a)$. For any two numbers $k, l \in \mathbb{N}$, we define the ergodic (k,l)–cross-sectional moment as

$$\mathbb{E}[x^k a^l] \equiv \int_x \int_a x^k a^l \, dF(x,a) \quad \forall k,l \in \mathbb{N}, \quad with \quad \mathbb{E}[x] = 0. \tag{19}$$

Aggregate dynamics. Fix an initial distribution of the state $F_0(x,a) = \mathbb{E}\left[\mathbb{1}_{\{(x_0,a_0)\leq(x,a)\}}\right]$. We define the impulse response function (IRF) for the m-th moment of the capital gap distribution under the initial distribution F_0, or $\text{IRF}_{m,t}(F_0)$, as the difference between its time t value and its ergodic value:

$$\text{IRF}_{m,t}(F_0,t) \equiv \underbrace{\mathbb{E}[x_t^m]}_{\text{transition}} - \underbrace{\mathbb{E}[x^m]}_{\text{steady-state}}. \tag{20}$$

Following Álvarez, Le Bihan and Lippi (2016), we define the cumulative impulse response (CIR), denoted by $\text{CIR}_m(F_0)$, as the area under the $\text{IRF}_{m,t}(F_0)$ curve across all dates $t \in (0,\infty)$:

$$\text{CIR}_m(F_0) \equiv \int_0^\infty \text{IRF}_{m,t}(F_0,t) \, dt. \tag{21}$$
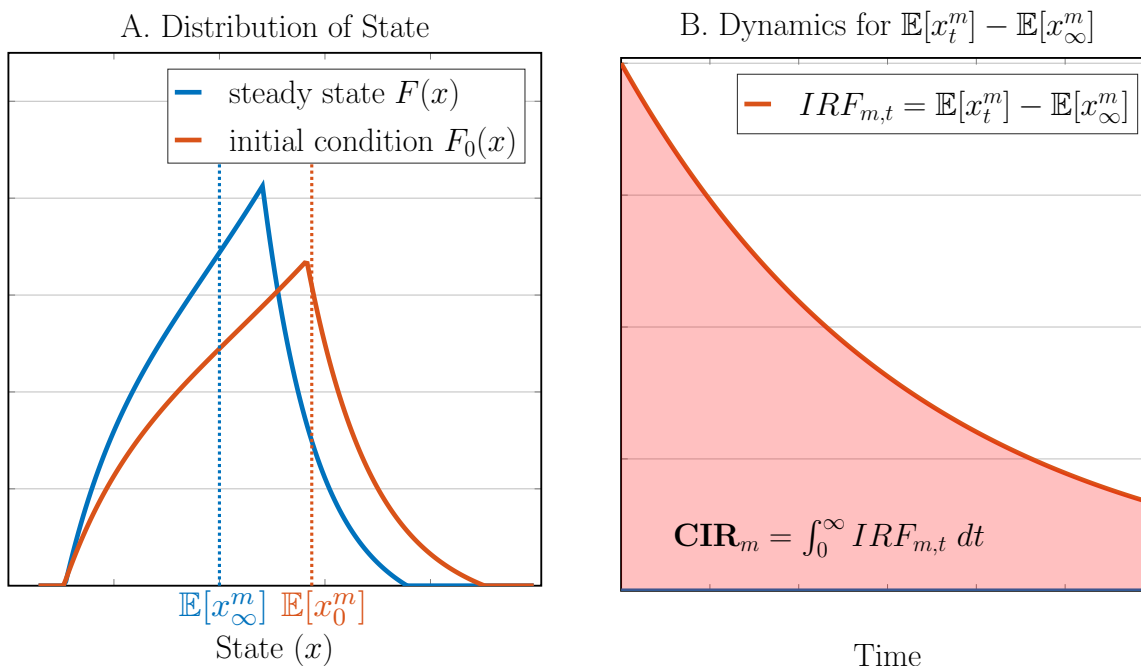
Figure I illustrates these two objects. In the left panel, we plot an initial marginal distribution $F_0(x)$ and the steady-state distribution $F(x)$; we also highlight the m-th moment of capital gaps $\mathbb{E}[x_0^m]$, which will be tracked on its way toward steady-state.[13] In the right panel, the solid line represents the impulse response of $\mathbb{E}[x_t^m]$, a function of time, and the area underneath it is the CIR. The CIR is our key measure of the convergence speed toward steady-state. The smaller the CIR, the faster the convergence.

Initial distribution as δ-perturbations around steady-state. For expositional purposes, we interpret the initial distribution F_0 as a small perturbation of the steady-state distribution that can be parsimoniously described in terms of one parameter δ. As a baseline, we focus on a particular perturbation consisting of a small uniform horizontal displacement of the distribution of capital gaps relative to the invariant F, i.e., a shock to the first moment of the distribution. This perturbation is labeled the "marginal response function" by Borovička, Hansen and Scheinkman (2014). If $f(x-\delta,a)$ denotes the initial density of capital gaps and $f_x(x,a)$ denotes its derivative with respect to x, we can approximate it as $f(x-\delta,a) \approx f(x,a) - \delta f_x(x,a)$. For $\delta > 0$, we observe that the initial density is equal to a right shift of the steady-state density. Afterward, the distribution will evolve according to the agents' policies and will converge back to its steady-state. Under this interpretation, we denote the $\text{CIR}_m(F_0)$ as $\text{CIR}_m(\delta)$.[14]

---

[13]Abusing notation, we denote the marginal distributions as $F(x)$ and $F(a)$.

[14]In Section 5.2 we consider more general perturbations around steady-state, such as mean-preserving spreads that capture aggregate second-moment shocks.

Figure I – Distribution Dynamics and Cumulative Impulse Response (CIR)

A. Distribution of State

- steady state $F(x)$
- initial condition $F_0(x)$

$\mathbb{E}[x_\infty^m]$ $\mathbb{E}[x_0^m]$

State (x)

B. Dynamics for $\mathbb{E}[x_t^m] - \mathbb{E}[x_\infty^m]$

- $IRF_{m,t} = \mathbb{E}[x_t^m] - \mathbb{E}[x_\infty^m]$

$\mathbf{CIR}_m = \int_0^\infty IRF_{m,t}\, dt$

Time

Notes: Panel A shows the steady-state distribution of the idiosyncratic state $F(x)$ and an initial distribution $F_0(x)$. It also illustrates the cross-sectional m-th moment to be tracked, from its initial value $\mathbb{E}[x_0^m]$ toward its steady-state $\mathbb{E}[x_\infty^m]$. Panel B shows the dynamics of the m-th moment in two ways: the IRF (solid line) and the CIR (area under the IRF).

3 Theoretical Results

We establish the theoretical relationships between observed data on adjustments $(\Delta x, \tau)$, steady-state moments and parameters, and aggregate dynamics measured via the CIR.

3.1 Structural link between micro data and steady-state moments

The first link connects the ergodic cross-sectional moments, the structural parameters of the stochastic process, and the reset state to the distribution of capital gaps Δx and adjustment dates τ obtained from a panel of observations. The relevance of this result is that, in many applications, the state x is likely to be unobservable, but the adjustment sizes Δx and durations τ are observable. This is the case in the investment model, as capital gaps are hard to observe but investment rates are readily available.

Proposition 1 derives a set of relationships between objects that we do observe—the distribution of investment and duration—with objects that we do not observe—the distribution of capital gaps.

Proposition 1. *Let $(\Delta x, \tau)$ be a panel of observations of adjustment size and inaction duration. Denote with $\mathbb{E}[\cdot], \mathbb{C}ov[\cdot]$, and $\mathbb{CV}^2[\cdot] = \mathbb{V}[\cdot]/\mathbb{E}[\cdot]^2$, respectively, the cross-sectional average, the covariance, and the coefficient of variation squared, conditional on adjustment. Then, the following expressions hold:*

1. *The reset capital gap is given by*

$$\hat{x} \;\; = \;\; \frac{\mathbb{E}[\Delta x]}{2} \left(1 - \mathbb{CV}^2[\tau] \right) + \frac{\mathbb{C}ov[\tau, \Delta x]}{\mathbb{E}[\tau]}. \tag{22}$$

2. *The drift and volatility of the capital gap process are recovered as*

$$\nu \;\; = \;\; -\frac{\mathbb{E}[\Delta x]}{\mathbb{E}[\tau]} \;\; ; \;\; \sigma^2 \;\; = \;\; \frac{\mathbb{E}[\Delta x^2]}{\mathbb{E}[\tau]} + 2\nu\hat{x}. \tag{23}$$

3. *Average age (time elapsed since last adjustment) relates to inaction duration as*

$$\mathbb{E}[a] \;\; = \;\; \frac{\mathbb{E}[\tau]}{2} \left(1 + \mathbb{CV}^2[\tau] \right) \tag{24}$$

4. *With non-zero drift ($\nu \neq 0$), the steady-state moments for any $m \geq 1$ are given by*

$$\mathbb{E}[x^m] \;\; = \;\; \frac{1}{m+1} \left\{ \frac{\hat{x}^{m+1} - \mathbb{E}\left[(\hat{x} - \Delta x)^{m+1} \right]}{\mathbb{E}[\Delta x]} - \frac{\sigma^2}{2\nu} m(m+1)\mathbb{E}[x^{m-1}] \right\}, \tag{25}$$

$$\mathbb{E}[x^m a] \;\; = \;\; \frac{1}{m+1} \left\{ \frac{\mathbb{E}[\tau]\mathbb{E}[x^{m+1}] - \mathbb{E}\left[\tau \, (\hat{x} - \Delta x)^{m+1} \right]}{\mathbb{E}[\Delta x]} - \frac{\sigma^2}{2\nu} m(m+1)\mathbb{E}[x^{m-1}a] \right\}. \tag{26}$$

5. *With zero drift ($\nu = 0$), the steady-state moments for any $m \geq 1$ are given by*

$$\mathbb{E}[x^m] \;\; = \;\; \frac{2}{(m+1)(m+2)} \frac{\mathbb{E}\left[(\hat{x} - \Delta x)^{m+2} - \hat{x}^{m+2} \right]}{\mathbb{E}[\Delta x^2]}, \tag{27}$$

$$\mathbb{E}[x^m a] \;\; = \;\; \frac{2}{(m+1)(m+2)} \left[\frac{\mathbb{E}\left[\tau(\hat{x} - \Delta x)^{m+2} \right]}{\mathbb{E}[\Delta x^2]} - \frac{\mathbb{E}[x^{m+2}]}{\sigma^2} \right]. \tag{28}$$

The proof is in the Appendix; it enumerates the formal assumptions for this proposition to hold. To show this result, we use three tools: Itō's Lemma, the Optional Sampling Theorem (OST), and the equivalence between the cross-sectional distribution of agents and the occupancy measure.[15] To understand the logic of the proof, assume there is a single state x. Apply Itō's Lemma to x_t^{m+1} with the initial condition $\hat{x}$ (right after adjustment), integrate between 0 and τ, and use $x_\tau \;=\; \hat{x} - \Delta x_\tau$ to obtain

$$\underbrace{(\hat{x} - \Delta x_\tau)^{m+1} - \hat{x}^{m+1}}_{\text{investment (observable)}} = \underbrace{\int_0^\tau \left[\nu(m+1)x_t^m + \frac{\sigma^2}{2}(m+1)m x_t^{m-1} \right] dt}_{\text{history of capital gaps (unobservable)}} + \underbrace{\sigma(m+1) \int_0^\tau x_t^m dW_t}_{noise}. \tag{29}$$

Equation (29) shows that the distribution of investment in the LHS is related to the history of capital gaps plus a noise term in the RHS. While we cannot recover each individual history, we can recover the *average* history between adjustments.

For this, take the expectation on both sides of (29) and observe that the noise term is a martingale with expectation zero by the OST, and obtain

$$\underbrace{\mathbb{E}\left[(\hat{x} - \Delta x_\tau)^{m+1} - \hat{x}^{m+1}\right]}_{\text{moments of investment}} = \underbrace{\nu(m+1)\mathbb{E}\left[\int_0^\tau x_t^m dt\right] + \frac{\sigma^2}{2}m(m+1)\mathbb{E}\left[\int_0^\tau x_t^{m-1} dt\right]}_{\text{average capital gap during inaction}}. \qquad (30)$$

The final step in relating the distributions of investment and capital gaps uses the occupancy measure. Intuitively, the average time a single agent's state spends at a given value is proportional to the number of agents with a state equal to that same value, where the constant of proportionality is the expected time between adjustments $\mathbb{E}[\tau]$.[16] Therefore, instead of measuring the average capital gap between adjustments for an agent, we can measure the average capital gap across agents:

$$\underbrace{\mathbb{E}\left[(\hat{x} - \Delta x_\tau)^{m+1} - \hat{x}^{m+1}\right]/\mathbb{E}[\tau]}_{\text{moments of investment and duration}} = \underbrace{\nu(m+1)\mathbb{E}\left[x^m\right] + \frac{\sigma^2}{2}m(m+1)\mathbb{E}\left[x^{m-1}\right]}_{\text{average capital gap across agents}}. \qquad (31)$$

Now we provide the economics behind each expression in Proposition 1.

Reset state. Equation (22) shows how to recover the reset state $\hat{x}$ from the micro data; this expression is derived from the restriction imposed by the normalization of the ergodic mean of the centralized distribution to zero. It has two components that jointly reflect how the reset state compensates for drift, for asymmetry in state-dependent policies, or a combination of both, ensuring that $\mathbb{E}[x] = 0$. To illustrate the compensation for drift, consider the family of fully time-dependent costs, which by construction do not exhibit asymmetric policies. In such models, $\mathbb{Cov}[\tau, \Delta x] = -\nu \mathbb{Var}[\tau]$ and equation (22) collapses to $\hat{x} = -\nu \mathbb{E}[a]$, so that the reset state compensates the average accumulated drift between adjustments, centralizing the ergodic mean at zero.[17]

To illustrate the compensation for asymmetric policies, consider a driftless state and a fully state-dependent cost in which the width of the lower and upper inaction triggers relative to the reset point are $|\hat{x} - \underline{x}|$ and $|\overline{x} - \hat{x}|$, respectively. Under no drift, the reset state is $\hat{x} = \mathbb{Cov}[\tau, \Delta x]/\mathbb{E}[\tau]$. Panel A of Figure II plots three distributions of capital gaps for different types of policies (these distributions may not be available to the econometrician). First, symmetric policies (in green) necessarily imply $\hat{x} = 0$ in order to obtain a zero ergodic mean. Our formula obtains the same conclusion directly from the data on adjustments—in this case $Pr(\Delta x = z|\tau) = Pr(\Delta x = -z|\tau) = 1/2$—and therefore the covariance between τ and Δx is zero. Now consider an asymmetric inaction region such that the upper trigger is closer to $\hat{x}$ than the lower trigger; for example, $4z = |\hat{x} - \underline{x}| > |\overline{x} - \hat{x}| = z$ for $z > 0$ (in red). In this case, the capital gap distribution is left-skewed and $\hat{x} = z$. Due to the asymmetry in the policy, the longer the duration of inaction, the higher the probability of a large positive investment; therefore, the formula identifies a positive reset state $\hat{x} = \mathbb{Cov}[\tau, \Delta x]/\mathbb{E}[\tau] = z > 0$. Analogously, the reset state is negative for right-skewed distributions (in blue).

---

[16]Formally, $\mathbb{E}\left[\int_0^\tau x_t^m dt\right] = \mathbb{E}[\tau]\mathbb{E}[x^m]$ for any m.

[17]Proof: $\mathbb{Cov}[\tau, \Delta x] = \mathbb{E}[\tau \Delta x] - \mathbb{E}[\tau]\mathbb{E}[\Delta x] = \mathbb{E}\left[\tau(-\nu\tau - \sigma\int_0^\tau W_t)\right] + \nu\mathbb{E}[\tau]^2 = -\nu\mathbb{E}[\tau^2] + \nu\mathbb{E}[\tau]^2 = -\nu\mathbb{Var}[\tau]$, where we have used the OST to kill the martingale $\mathbb{E}\left[\tau\int_0^\tau W_t\right] = 0$.

Figure II – Reset state for asymmetric policies in fully state-dependent models

A. Fully state dependent

B. CalvoPlus
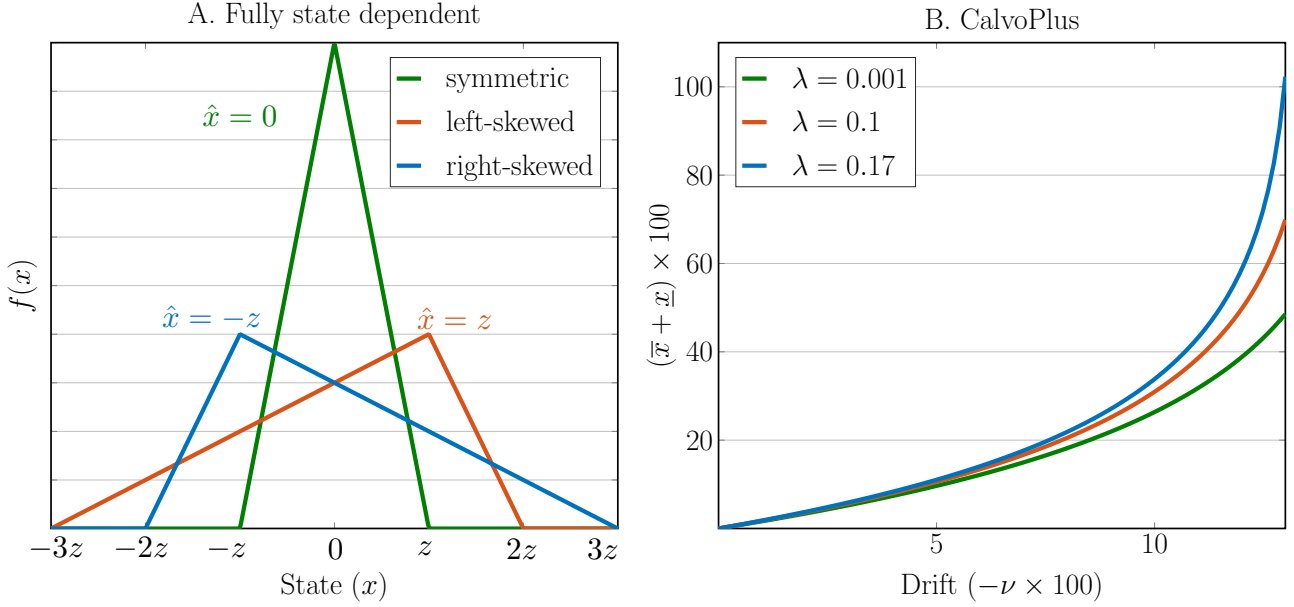
Notes: Panel A describes the distribution of state x in a fully state-dependent model. The symmetric distribution (green line) has widths of $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (z, z)$, the left-skewed distribution (red line) $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (4z, z)$, and right-skewed distribution (blue line) $(|\hat{x} - \underline{x}|, |\overline{x} - \hat{x}|) = (z, 4z)$. The reset states are 0, z, and $-z$, respectively. Panel B shows the levels of asymmetry in policy and drift that together imply a zero reset state $\hat{x} = 0$, for fixed parameters $(\sigma, \underline{x}) = (0.275, -0.49)$ and three values of $\lambda \in \{0.001, 0.1, 0.17\}$.

Lastly, regarding the interactions, policy asymmetry may dampen or amplify the drift's effect in the reset state. For illustration, consider the Bernoulli fixed-cost model and fix a set of parameters. Panel B in Figure II shows the combination of values for $(-\nu \times 100, \overline{x} + \underline{x})$ such that the reset state is zero $\hat{x} = 0$. When $\nu = 0$, only symmetric policies generate a zero reset $\hat{x}$. As the drift increases, the upper limit $\overline{x}$ increases as well to compensate for the drift. The covariance term informs this interaction, as in this case $\hat{x} = -\nu(\mathbb{E}[\tau] - \mathbb{E}[a]) + \mathbb{C}ov[\tau, \Delta x]/\mathbb{E}[\tau]$. Note that the asymmetry is also increasing in the parameter λ (the arrival of free adjustment opportunities): As λ increases, the fraction of state-dependent (asymmetric) investments decreases, and it must be compensated for by a more asymmetric policy.

Drift and volatility of stochastic process. Expressions in (23) provide a guide to infer the parameters of the stochastic process. The first expression shows how to infer drift ν from the average investment rate in the data, scaled by the adjustment frequency. Intuitively, the average depreciation (plus common trend) has to be equal to the investment rate to have an ergodic distribution. The second expression shows how to infer volatility σ from the dispersion in investment rates, scaled by the frequency and corrected by the drift. With a zero drift, higher idiosyncratic volatility reduces the average duration or increases the average investment size; with non zero drift, investment's second moment also reflects the drift and (23) shows how to correct for it.[18]

---

[18]Álvarez, Le Bihan and Lippi (2016) obtain this expression for idiosyncratic volatility in the driftless case.

Average age. Equation (24) relates average age to the average and the dispersion in duration, measured through the coefficient of variation. The relationship with the average duration is straightforward. To understand why the duration dispersion affects average age, recall a basic property in renewal theory: The probability that a random firm has an expected duration of inaction of τ is increasing in τ, i.e., larger stopping times are more representative in the capital-gap distribution.[19] Therefore, dispersion in duration reflects that some firms take a long time to adjust, and on top of that, those firms are more representative in the economy; this raises the average age. There are two special cases in which the dispersion in duration does not matter for average age. With an exponential distributions of stopping times, the coefficient of variation is equal to 1 and (24) collapses to $\mathbb{E}[a] = \mathbb{E}[\tau]$. The second case is with a fixed adjustment date T, for which $\mathbb{E}[a] = T/2$. In the empirical application in Section 4, we measure both sides of expression (24) and verify that it is satisfied with a high degree of precision.

Ergodic moments of x. Equations (25) and (27) provide formulas to compute the ergodic moments of the gap x using observed changes in the gap Δx. To build the intuition, consider the symmetric and driftless case ($\nu = \hat{x} = 0$) and set $m = 2$. Then the formula for the second moment yields

$$\mathbb{V}[x] = \frac{\mathbb{E}\left[\Delta x^4\right]}{6\mathbb{E}\left[\Delta x^2\right]} = \frac{\mathbb{E}\left[\Delta x^2\right]\mathbb{CV}\left[\Delta x^2\right]}{6}. \tag{32}$$

This expression shows how the dispersion of capital gaps in the LHS relates to the average and the dispersion (the coefficient of variation) of the adjustment size Δx^2 in the RHS. If the data shows large adjustment size on average, it reflects that capital gaps drift far away from zero increasing the ergodic variance; this force is amplified if there is dispersion in adjustment sizes, as large adjustments are more representative of the population of capital gaps (the same renewal logic as applied to age above). For $\nu > 0$, the moments are different but the same idea applies.

Now consider the case $m = 3$, which relates to the asymmetry in the distribution of capital gaps. A natural, but incorrect, conclusion is that a right-skewed investment rate distribution implies a left-skewed capital gap distribution. This is not necessarily true. The reason is that if the probability of adjustment is low for positive capital gaps, then we should not see positive investment rates. They are more representative, however, since they have larger duration. The theory can measure this force by comparing the fourth moment of the distribution versus the third moment. If $\nu < 0$, then

$$Sign\left(Skew[x]\right) = Sign\left(\mathbb{E}[x^3]\right) = Sign\left(-1 + \frac{4\sigma^2}{|\nu|}\frac{\mathbb{E}[\Delta x^3]}{\mathbb{E}[\Delta x^4]}\right). \tag{33}$$

Therefore, the capital gap distribution features positive skewness as long as $\mathbb{E}[\Delta x^3]/\mathbb{E}[\Delta x^4] > |\nu|/4\sigma^2$.

Joint ergodic moments of x and age. Equations (26) and (28) provide formulas for the joint ergodic moments of the gap and its age. Let us focus on their covariance, as it is a key moment for the CIR_1 under all the models we study in the next section. Following equation (26) and assuming $m = 1$ and

---

[19]This property has been widely studied in labor economics when thinking about long-term unemployment. For example, Mankiw (2014)'s textbook *Principles of Macroeconomics* states: "Many spells of unemployment are short, but most weeks of unemployment are attributable to long-term unemployment."

$\nu < 0$, we have that the covariance is composed of a negative and a positive term:

$$\mathbb{C}ov[x, a] = -\frac{1}{2\nu} \left\{ \underbrace{\frac{\mathbb{E}[\Delta x^3]}{3\mathbb{E}[\Delta x]} + \sigma^2 \mathbb{E}[a]}_{>0} - \underbrace{\left(\mathbb{E}\left[\Delta x^2\right] + \frac{\mathbb{C}ov[\tau, \Delta x^2]}{\mathbb{E}[\tau]} \right)}_{<0} \right\}. \tag{34}$$

Therefore, the covariance could be positive or negative depending on two forces. The first force is related to the drift: Over time, firms' capital gaps depreciate and therefore we should expect a negative covariance between age and capital gaps. The second force, related to asymmetry in adjustment policies, pushes toward a positive covariance: If idiosyncratic shocks are large relative to the drift and the investment distribution is right-skewed (the probability of adjustment is lower for firms with capital gaps above the reset than for firms with capital gaps below the reset)—which could be the result, for example, of downward irreversibility—then the covariance is positive. The theory shows how to discipline these two forces with data on observables as shown in (34), where the term $\mathbb{E}[\Delta x^3]$ primarily captures the asymmetries while the term $\mathbb{C}ov[\tau, \Delta x^2]$ primarily captures the drift.

Application: Markups and misallocation. Let us explain a straightforward application of the results in Proposition 1. The expressions can be used to compute cross-sectional moments of the distribution of firm-level markups using pricing data without the need for data on marginal costs; we see this strategy as complementary to the so-called demand approach (Berry, Levinsohn and Pakes, 1995) or the more recent production approach (De Loecker and Warzynski, 2012; De Loecker and Eeckhout, 2017). Similarly, one can compute moments of firm-level marginal products of capital or labor using inputs' adjustment data, without additional requirements on the production technology (Hsieh and Klenow, 2009; David and Venkateswaran, 2019).

Another interesting application of the results for non-zero drift would be to revisit the role of positive inflation for price-setting and its welfare costs, as studied by Burstein and Hellwig (2008), Gagnon (2009), and Blanco (2019).

3.2 Structural link between steady-state moments and aggregate dynamics

This section shows that for the class of adjustment cost specifications described in Section 2.3, the CIR is a linear combination of ergodic moments, in which the coefficients depend on the stochastic process parameters and adjustment cost specification.

As a first step, Proposition 2 expresses the CIR in (21) as the solution to a collection of stopping-time problems, indexed by initial conditions. The result is surprising: It is only necessary to track each firm until its *first* adjustment, even in the presence of drift and asymmetric policies. The reason is that the first adjustment incorporates all deviations from steady-state due to the fact that the reset state is memoryless; any additional adjustments are driven purely by idiosyncratic conditions that produce the ergodic moment of interest and thus do not contribute to deviations from steady-state. Importantly, this result depends on the independence of prices with respect to the firms' distribution, which implies that the steady-state policies hold along the transition path.

We generalize the result in Álvarez, Le Bihan and Lippi (2016), who show that to compute the CIR of the first moment one only needs to keep track of an agent until its first adjustment in a driftless economy with symmetric policies, as the cross-sectional average of subsequent adjustments is exactly equal to zero. Here, we show that this property extends to any moment of the distribution $m > 1$, any arbitrary Markovian stopping policy, and any Markovian law of motion of the uncontrolled state. Subsequent adjustments do not cancel out to zero, but are equal to the steady-state moment.

Proposition 2. *The CIR_m can be written recursively as*

$$CIR_m(F_0) = \int v_m(x, a) dF_0(x, a), \tag{35}$$

where the value function for an agent with initial state (x, a) is given by

$$v_m(x, a) \equiv \mathbb{E}\left[\int_0^\tau (x_t^m - \mathbb{E}[x^m]) dt \Big| (x, a)\right]. \tag{36}$$

The idea behind the proof of Proposition 2 is to exchange the integral across agents (the cross-section) with the infinite time integral (the time series).[20] Then, we show that the *first adjustment* incorporates all deviations from steady-state and the additional adjustments equal the ergodic moment $\mathbb{E}[x^m]$. This implies that the value function $v_m(x, a)$ equals zero after the first adjustment. For that reason, the infinite limit in the time integral gets substituted for the first stopping date $t = \tau$. Now we proceed to characterize the CIR for different variations of the random fixed-cost model.

3.2.1 Bernoulli fixed-cost model

Consider the Bernoulli fixed-cost model (fixed cost with free opportunities of adjustment). Proposition 3 derives a simple expression for the CIR_m, as a linear combination of two simple moments of the joint distribution $F(x, a)$.

Proposition 3. *In the Bernoulli fixed-cost model, up to first order, the CIR_m is given by*

$$CIR_m(\delta)/\delta = \frac{\mathbb{E}\left[x^{m+1}\right] - \nu\mathbb{C}ov\left[x^m, a\right]}{\sigma^2} + o(\delta). \tag{37}$$

Equation (37) shows that up to first order, there exists a one–to–one mapping between two ergodic moments and the CIR. The logic behind the proof is that the CIR and the ergodic moments share the same finite basis $\Omega = \{e^{\xi_1 x}, e^{\xi_2 x}, e^{\xi_1 x} x, e^{\xi_2 x} x, \{x^i\}_{i=0}^m\}$, where ξ_1, ξ_2 are the characteristic roots of the HJB satisfied during periods of inaction. After this observation, the rest involves simple coefficient matching and algebra.

To build the intuition for this result, let us focus on the case $m = 1$. Consider $\nu = 0$ so that the CIR_1 is given exclusively by $\mathbb{V}ar[x]/\sigma^2$, the dispersion of capital gaps normalized by the idiosyncratic volatility. This dispersion encodes information about agents' *responsiveness* to idiosyncratic shocks (the higher the ratio, the less responsive the policy) and, in turn, the responsiveness determines the speed of convergence

---

[20]This can be done due to the ergodic properties of the problem and the fact that moments are finite.

to the steady-state. In the case $\nu \neq 0$, the covariance between capital gap age and the investment rate appears in the expression to correct for the dispersion generated by the drift (which is orthogonal to the dispersion due to idiosyncratic shocks). The concept of responsiveness is tightly linked to the selection effects, formalized in symmetric models by Elsby and Michaels (2019).

Now consider $m = 2$. The dynamics of the second moment can have either sign. Assume $\nu < 0$ and σ sufficiently large. Then the covariance is necessarily positive, i.e., $\mathbb{C}ov\left[x^2, a\right] > 0$, since larger age is associated with larger accumulated gaps. Then the sign of the CIR_2 depends on the asymmetry of the capital gap distribution, measured through its third moment $\mathbb{E}[x^3]$, which in turn depends on how much an asymmetric policy compensates for the drift (recall Figure II).

Generally, aggregate dynamics for the m-th moment are structurally linked to the $m + 1$ steady-state moment. These connections can be useful to think about the cyclicality in the second moment of investment, as documented by Bachmann and Bayer (2014), or the cyclicality in the skewness in sales growth as documented by Bloom, Guvenen and Salgado (2016). There are some papers that already point towards these relationships. For instance, Ehouarne, Kuehn and Schreindorfer (2016) analyze the role of the skewness of investment rates in order to generate misallocation cycles.

Revisiting a known case. Álvarez, Le Bihan and Lippi (2016) characterize the CIR for $m = 1$, zero drift, and a symmetric policy ($\hat{x} = 0$) and obtain their well-known 'kurtosis' formula. Their result is nested in our formulas. Combining the CIR_1 in the Bernoulli fixed-cost model (37) with the expression for σ^2 in (23) and the expression for $\mathbb{V}ar[x]$ in (27), we obtain the kurtosis formula for the CIR_1:

$$\underbrace{\text{CIR}_1(\delta)/\delta}_{\text{aggregate dynamics}} \approx \underbrace{\mathbb{V}ar[x]/\sigma^2}_{\text{responsiveness}} = \underbrace{\mathbb{E}[\tau]\mathbb{K}ur[\Delta x]/6}_{\text{microdata}}. \tag{38}$$

Note that the ratio of ergodic variance to shock volatility—the intermediate link between aggregate dynamics and the microdata—captures the relative importance of shocks vis-à-vis responsiveness, as put forward by Berger and Vavra (2019).

Discussion. There are many advantages of working within Bernoulli fixed costs. There is a simple closed-form expression for the CIR with an intuitive economic interpretation; the two sufficient statistics involved are easily computed from the data (using the results in Proposition 1), and it is quite general in the sense that it accommodates arbitrary inaction regions and random *iid* adjustments. The disadvantage is that it imposes a strong assumption for adjustment costs. Nevertheless, we find that it provides a good approximation of the CIR_m for more general cases, as shown in Section 4.5. As an intermediate step toward characterizing the general random fixed-cost model, we consider fully time-dependent models.

3.2.2 Fully time-dependent models

With fully time-dependent adjustment costs, the distribution of stopping times is independent of the capital gap. Proposition 4 characterizes the CIR_m for this class of models.

Proposition 4. *Let $\phi \equiv \sigma^2/2\nu$. In fully time-dependent models, up to first order, the CIR_m is given by*

$$CIR_m(\delta)/\delta \quad = \quad \sum_{k=1}^{m} \frac{m!}{k!} (-\phi)^{m-k} \Gamma_k \ + \ o(\delta). \tag{39}$$

$$\Gamma_k \quad = \quad k\mathbb{E}\left[x^{k-1}a\right] + \mathbb{1}_{\{k\geq 2\}}\phi k(k-1)\mathbb{E}\left[x^{k-2}a\right]. \tag{40}$$

We leave discussion of the proof for later, as it is a nested case of the random fixed-cost model explained below. To understand the economics behind this relationship, consider first the case $m = 1$:

$$\text{CIR}_1(\delta)/\delta \quad = \quad \mathbb{E}\left[a\right] \ + \ o(\delta). \tag{41}$$

Average age provides information about the speed at which the average firm adjusts to the perturbation from the steady-state. Consider a frictionless limit in which all firms continuously invest to bring capital gaps to zero. Since capital in all firms would have age equal to zero, the economy reaches its steady-state immediately. Notice that in the presence of drift and asymmetric policies, the sufficient statistic for fully time-dependent models $\mathbb{E}[a]$ differs to that of the Bernoulli fixed-cost model $(\mathbb{V}ar\left[x\right] - \nu\mathbb{C}ov\left[x,a\right])/\sigma^2$. However, with zero drift and symmetric policies, one can show that $\mathbb{E}[a] = \mathbb{V}ar[x]/\sigma^2$, implying that the sufficient statistic for both models coincides.[21] As shown by Alvarez, Lippi and Passadore (2017), this is true as long as shocks δ are small.

Now consider the case $m = 2$, where we have that[22]

$$\text{CIR}_2(\delta)/2\delta \quad = \quad \mathbb{C}ov\left[x,a\right] \ + \ o(\delta). \tag{42}$$

From this expression, it is clear that a positive δ-perturbation always generates a negative second-moment CIR whenever $-\nu < 0$. The value of the covariance depends on the model. Let us illustrate this point with an example. Computing the covariance in terms of moments of the duration distribution,[23] we can derive expressions for a fixed date and *iid* exponentially distributed dates: $\text{CIR}_2^{fixed}(\delta)/\delta = \nu\mathbb{E}[\tau]\left(1 + 2\mathbb{E}[\tau]/3\right)$ and $\text{CIR}_2^{iid}(\delta)/\delta = 2\nu\mathbb{E}[\tau]\left(1 + 2\mathbb{E}[\tau]\right)$, which are negative in both cases; therefore, the second moment shrinks in the transition. Note that, for the same average duration, the exponential date model has a larger covariance in absolute value than with a fixed date, and thus the shrinkage is larger. The underlying reason is due to differences in the tail of the duration distribution generated by each model.

---

[21]See Online Appendix B.1 for the proof.

[22]The formula implies $\text{CIR}_2(\delta)/\delta = \Gamma_2 - 2\phi\Gamma_1 = 2\mathbb{C}ov\left[x,a\right] + 2\phi\mathbb{E}[a] - 2\phi\mathbb{E}[a] \ + \ o(\delta) = 2\mathbb{C}ov\left[x,a\right] \ + \ o(\delta)$.

[23]The expression reads $\mathbb{C}ov[x,a] = \frac{\nu}{\mathbb{E}[\tau]}\left(\frac{\mathbb{E}[\tau^2]}{2} + \frac{\mathbb{E}[\tau^3]}{3}\right)$. With fixed dates we substitute $\mathbb{E}[\tau^m] = \mathbb{E}[\tau]^m$; with *iid* exponential dates we substitute $\mathbb{E}[\tau^m] = m!\mathbb{E}[\tau]^m$.

3.2.3 General random fixed-cost model

Now we proceed to characterize aggregate dynamics under very general random fixed costs. The strategy is as follows. We set as an upper bound the value of the CIR obtained for fully time-dependent models. This value reflects adjustments along the *intensive margin* exclusively, as agents cannot affect the duration of inaction. Then, we consider an additional term that captures adjustments through the *extensive margin*, or changes in the duration—which are available in models with a state-dependent component—that can only reduce the CIR.

Two challenges are involved in bringing discipline to the extensive margin. First, the extensive margin not only depends on the immediate response of the aggregate duration but also reflects all current and future changes in duration. Second, even if we had the whole sequence of adjustments in duration that follows a perturbation, the extensive margin also depends on the capital gaps of the particular set of firms selected to invest. We propose a way to bring discipline to the extensive margin in a general setting. See Caballero and Engel (2007) for a discussion on the importance of this margin in a static setting.

How to characterize the extensive margin? In order to characterize the extensive margin, we introduce the following auxiliary function:

$$g_m(x) \equiv \mathbb{E}[x_\tau^m | \hat{x} + x] - \mathbb{E}^{\hat{x}}[(x_\tau + x)^m | \hat{x}]. \tag{43}$$

The first term in (43) equals the expected capital gap at the moment of adjustment when the initial condition is $\hat{x} + x$; the second term equals the expected capital gap at the moment of adjustment plus a deterministic increase of size x when the initial condition is $\hat{x}$. The difference between these two functions of x provides information on how the stopping-time policy depends on the initial condition and how it correlates with the capital gap. To see this clearly, recall that the expected capital gap at the moment of adjustment is equal to $x_\tau = \hat{x} - \Delta x$ and we can re-express $g_m(x)$ in the following way:

$$g_m(x) = \mathbb{E}\left[(\hat{x} + x - \nu\tau^{\hat{x}+x} - \sigma W_{\tau^{\hat{x}+x}})^m\right] - \mathbb{E}\left[(\hat{x} + x - \nu\tau^{\hat{x}} - \sigma W_{\tau^{\hat{x}}})^m\right], \tag{44}$$

where τ^z is the stopping time with initial condition z. In equation (44) we observe that if τ is independent of the initial condition, i.e., $\tau^{\hat{x}+x} = \tau^{\hat{x}} = \tau$, as in time-dependent models, then $g_m(x) = 0$ for all x, implying that the extensive margin is null. For other models, the derivatives of g_m with respect to the initial condition, evaluated at zero, provide a micro-elasticity of firms' idiosyncratic response to the new initial conditions though changes in its stopping time τ. With this function at hand, we proceed to characterize the CIR_m for the random fixed-cost model.

Proposition 5. *Let $\phi \equiv \sigma^2/2\nu$ and define the auxiliary function $g_m(x) \equiv \mathbb{E}[x_\tau^m | \hat{x} + x] - \mathbb{E}^{\hat{x}}[(x_\tau + x)^m | \hat{x}]$. In the generalized hazard model, for every $m \geq 1$, the CIR_m is given by*

$$CIR_m(\delta)/\delta = \sum_{k=1}^m \frac{m!}{k!}(-\phi)^{m-k}(\Gamma_k + \Psi_k) - \mathbb{E}[x^m]\Psi_0 + o(\delta), \tag{45}$$

21

where the margins of adjustment are given by

$$(intensive) \quad \Gamma_k = k\mathbb{E}[x^{k-1}a] + \mathbb{1}_{\{k\geq 2\}}k(k-1)\phi\mathbb{E}[x^{k-2}a], \quad (46)$$

$$(extensive) \quad \Psi_k = \sum_{j=0}^{\infty} \psi_{k,j}\mathbb{E}[x^j], \quad (47)$$

$$\psi_{k,j} = \frac{1}{\nu}\sum_{i\geq j}^{\infty}\frac{\hat{x}^{i-j}}{i!j!}\left[\frac{d^{i+1}g_{k+1}(0)}{dx^{i+1}}/(k+1) - \frac{d^i g_k(0)}{dx^i}\right]. \quad (48)$$

The total effect of the δ-perturbation is an area with height δ and a base given by the recursive sum of two components: Γ_k, which measures adjustments through the intensive margin and its expression is identical to that for time-dependent models in equation (41); and Ψ_k, which measures adjustments through the extensive margin and is expressed as a linear combination of ergodic moments. In turn, the weights $\psi_{m,j}$, or micro-elasticities, are written in terms of derivatives of the auxiliary function g_m, evaluated at zero.

The proof of Proposition 5 is constructive and has two steps. The first step constructs two Bellman equations of a representative agent to characterize the intensive Γ_m and the extensive Ψ_m margins of adjustment. The history independence in the stochastic processes and policies allows us to collapse all of the ex post heterogeneity due to idiosyncratic shocks and frictions to the problem of a representative agent. Importantly, this aggregation result *does not imply* that heterogeneity is irrelevant for aggregate dynamics; it says that all heterogeneity can be summarized in a compact way. The second step proceeds similarly to Proposition 1 and expresses each Bellman equation as a function of ergodic moments using a combination of Itō's Lemma, Optional Sampling Theorem, and the occupancy measure.

Propositions 3 and 4 are subcases of Proposition 5. In fully time-dependent models, since the duration adjustment is independent of the state, we have that $\psi_{k,j} = 0$ and $\Psi_k = 0$ for all k, and therefore CIR_m collapses to (41). The Bernoulli fixed-cost model, due to its constant hazard of adjustment, allows for a closed-form solution for $\psi_{k,j}$. However, computations are quite involved. Thus we proved Proposition 3 in a simpler way using principles of linear algebra, as outlined above.

Constructing micro-elasticities from the data. To construct micro-elasticities, we need to ask: What piece of $g_m(x)$ is observable in the data and what is not? The object $\mathbb{E}\left[(\hat{x} - \Delta x + x)^m\right]$ is an observable statistic, as it depends on the steady-state investment rates. The objects $\mathbb{E}[x^j]$ and $\hat{x}$ can also be recovered from the data, as we have shown before. Therefore, the only object that might not be directly observable is $\mathbb{E}^{\hat{x}+x}\left[(\hat{x} - \Delta x)^m\right]$, which measures the elasticity of investment with respect to changes in initial conditions. Guided by the theory, we suggest that this elasticity is the key object that future research should focus on computing, both in the data and in the models.

Two papers use an adequate methodology and data to construct micro-elasticities. First, in the pricing literature, Karadi and Reiff (2019) study the immediate monthly price response to a change in the VAT using the Hungarian CPI. Since a change in the VAT proxies a cost-push shock, the experiment is equivalent to an increase in a firm's markup in the same proportion; this design would allow one to compute the micro-elasticity of the expected price change to initial conditions. Second, in the investment

literature, Zwick and Mahon (2017) exploit shifts in accelerated depreciation to estimate the effect of temporary tax incentives on equipment investment; such a design would allow one to compute the micro-elasticity of expected investment to initial conditions.

In the following section we apply our theoretical results to the data.

4 Application: Investment Dynamics

In this section, we revisit the investment model from Section 2 and apply our tools using establishment-level data from Chile. In the process, we assess the nature of adjustment costs, the magnitude of capital misallocation in steady-state, and the aggregate dynamics of aggregate capital.

4.1 Data description

Sources. We use yearly micro data on the cross-section of manufacturing plants in Chile from the Annual National Manufacturing Survey (*Encuesta Nacional Industrial Anual*) for the period 1979 to 2011.[24] Information on depreciation rates and price deflators used to construct the capital series comes from National Accounts and Penn World Tables. The Online Data Appendix presents all of the details on the data, construction of variables, and detailed analysis of each capital category separately: structures, vehicles, machinery, and equipment. Here we focus on the total capital stock and structures, a category that represents approximately 30% of all investment in the manufacturing sector and features the strongest lumpy behavior. We consider all plants that appear in the sample for at least 10 years (more than 60% of the sample) and have more than 10 workers.

Capital stock and investment rates. We construct the capital stock series through the perpetual inventory method (PIM).[25] Let a plant's ω stock of capital on year t be given by

$$K_{\omega,t} = (1 - \delta_j)K_{\omega,t-1} + I_{\omega,t}/D_t \quad \text{for } K_{\omega,t_0} \text{ given,} \tag{49}$$

where depreciation rate δ is a type-specific time-invariant depreciation rate; price deflators D_t are gross fixed capital formation deflators by capital type; and initial capital K_{i,t_0} is given by plants' self-reported nominal stock of capital by type at current prices for the first year in which they report a non negative capital stock. Gross nominal investment $I_{\omega,t}$ is based on information on purchases, reforms, improvements, and sales of fixed assets reported by each plant in the survey:

$$I_{\omega,t} = puchases_{\omega,t} + reforms_{\omega,t} + improvements_{\omega,t} - sales_{\omega,t}. \tag{50}$$

---

[24]These data have been used by Liu (1993) to examine the role of turnover and learning on productivity growth; by Tybout (2000) to survey the state of the manufacturing sector in developing economies; and more recently by Oberfield (2013) to study productivity and misallocation during crises.

[25]See Section A.2 for details on the PIM method and several checks on the data.

Once we construct the investment and capital stock series, we define investment rate $i_{\omega,t}$ as the ratio of real gross investment to capital stock:[26]

$$i_{\omega,t} \equiv \frac{I_{\omega,t}/D_t}{K_{\omega,t}}. \tag{51}$$

4.2 Construction of capital gaps and duration

To apply the theory, for each plant ω and each inaction spell k we record the capital gap change upon action $\Delta x_{\omega,k}$ and the spell's duration $\tau_{\omega,k}$. Using the information on investment rates, we construct capital gap changes as[27]

$$\Delta x_{\omega,k} = \begin{cases} \log\left(1 + i_{\omega,k}\right) & \text{if} & |i_{\omega,k}| > \underline{i} \\ 0 & \text{if} & |i_{\omega,k}| < \underline{i}, \end{cases} \tag{52}$$

where $\underline{i} > 0$ is a parameter that captures the idea that small maintenance investments do not incur the fixed cost. Following Cooper and Haltiwanger (2006), we set $\underline{i} = 0.01$, such that all investments smaller than 1% in absolute value are excluded and considered as inaction. Additionally, we truncate the distribution at the 1st and 99th percentiles. Finally, we compute a spell's duration as the difference between two adjacent adjustment dates:

$$\tau_{\omega,k} = T_{\omega,k} - T_{\omega,k-1}. \tag{53}$$

Figure III plots the cross-sectional distribution of capital gap changes. Following the investment distribution, it presents sizeable asymmetry and a large positive skewness. In each figure, we show the distribution for two subsamples: observations with spell duration above the average (dark bars) and spell duration below the average (white bars). The two distributions lie on top of one another, a sign of lack of covariance between adjustment size and duration.

As the next step, we put the theory to work by computing the cross-sectional statistics of capital gap changes and duration in order to back out the parameters of the stochastic process as well as the ergodic moments.

4.3 Putting the theory to work

The relationships derived in Proposition 1 tell us how to use cross-sectional data on capital gaps and duration to pin down the parameters of the productivity process, the reset point, and the ergodic moments, which in turn are used to construct the CIR. Table I summarizes the statistics calculated from the micro data, which serve as inputs into the formulas, as well as the theory's output.[28]

---

[26]Table X in the Online Appendix presents descriptive statistics on investment rates by capital category. For comparison, the table also includes the numbers reported by Cooper and Haltiwanger (2006) for 7,000 US manufacturing plants between 1972 and 1988 and by Zwick and Mahon (2017) from tax records for US firms.

[27]In contrast to equation (11), which computes investment as the difference in the capital stock between two instants, for the data we consider investment as the difference between two consecutive years.

[28]Section 4.6 explains how we address heterogeneity by industrial sector and plant size, and time trends, among other

Figure III – Distribution of capital gap changes by duration
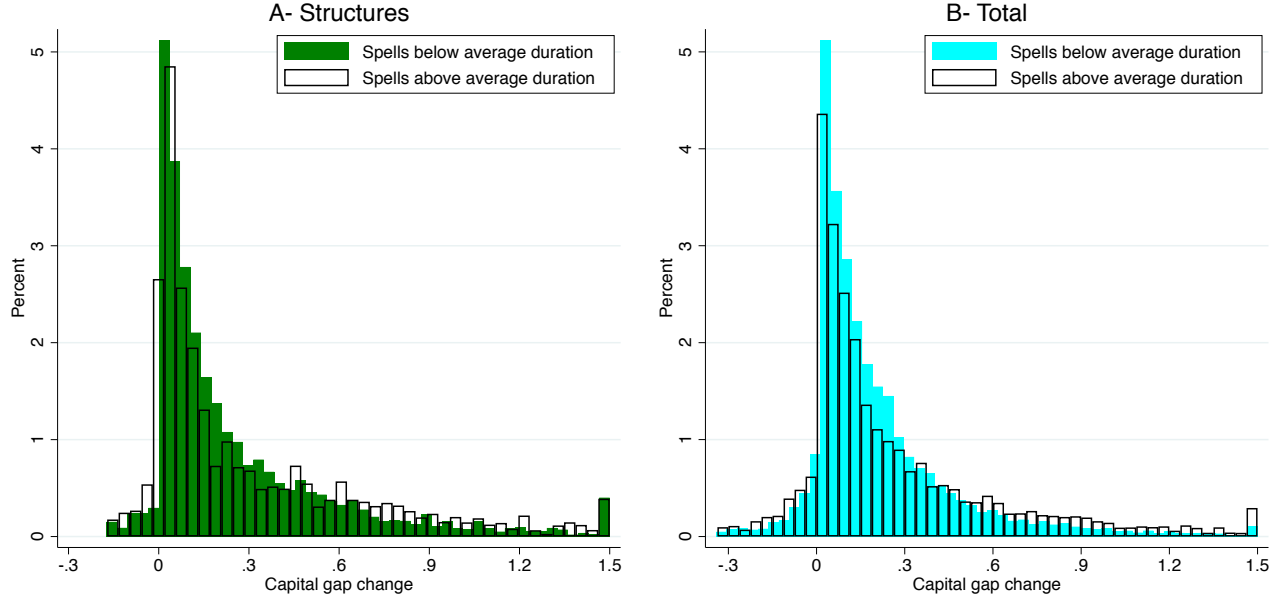
Notes: Panel A plots the distribution of capital gap changes Δx in structures, while Panel B plots the distribution for total capital. Solid bars are the distribution conditional on inaction spells with a duration below the average; white bars are the distribution conditional on inaction spells with a duration above the average. We consider a balanced panel with a truncation of 1st and 99th percentiles, depreciation rates reported in Table II in the Data Appendix, and an inaction threshold of $\underline{i} = 0.01$.

Table I – Inputs from Micro Data and Outputs from the Theory

| | Inputs from Data | | | Outputs from Theory | |
	Structures	Total		Structures	Total
Frequency			**Parameters**		
$\mathbb{E}[\tau]$	2.441	1.711	ν	-0.111	-0.129
$\mathbb{CV}^2[\tau]$	1.096	0.847	σ^2	0.075	0.067
			$\hat{x}$	0.013	0.041
Capital Gaps			**Steady-state Moments**		
$\mathbb{E}[\Delta x]$	0.271	0.221	$\mathbb{V}ar[x]$	0.226	0.148
$\mathbb{E}[\Delta x^2]$	0.191	0.132	$\mathbb{E}[a]$	2.558	1.581
$\mathbb{E}[(\hat{x} - \Delta x)^3]$	-0.184	-0.098	$\mathbb{C}ov[a, x]$	0.900	0.406
Covariances			**Aggregate Dynamics via CIR$_1$**		
$\mathbb{C}ov[\tau, \Delta x]$	0.064	0.042	Bernoulli fixed cost	4.328	3.013
$\mathbb{E}[\tau(\hat{x} - \Delta x)^2]$	0.535	0.253	Time dependent	2.558	1.581

Notes: Authors' calculations using establishment-level survey data for Chile. We consider a balanced panel with a truncation of 1st and 99th percentiles, depreciation rates reported in Table II in the Data Appendix, and an inaction threshold of $\underline{i} = 0.01$.

The left part of Table I shows the inputs from the data: cross-sectional statistics for frequency, capital gaps, and covariances between them. The right part of the table shows the outputs from our theory: parameters $(\nu, \sigma^2, \hat{x})$ and ergodic moments $(\mathbb{E}[a], \mathbb{V}ar[x], \mathbb{C}ov[x,a])$. For aggregate dynamics via the CIR$_1$, we show the values obtained for our formulas in the Bernoulli fixed-cost model in (37) and time-dependent model in (41). We do not report the numbers obtained numerically for the general random fixed-cost model, but these are very close to the Bernoulli case. We take this as a confirmation that the formula with variance and covariance holds approximately in more general settings.

Inputs from micro data. Consider first the distribution of expected times τ. We obtain an average expected time to adjustment for structures of $\mathbb{E}[\tau] = 2.4$ years with a large dispersion, suggesting substantial heterogeneity in adjustment times across plants. Now consider the distribution of capital gaps; it has an average of $\mathbb{E}[\Delta x] = 0.27$ and a second moment of $\mathbb{E}[\Delta x^2] = 0.19$, and it is right-skewed. The covariance between adjustment size and expected time is almost zero $\mathbb{C}ov[\tau, \Delta x] = 0.06$, as suggested by Figure III.

Outputs from theory: Parameters. We will now explain the parameter values implied by our formulas. From (23), the implied drift, which captures the depreciation rate, productivity growth, and changes in relative prices, equals

$$\nu = -\frac{\mathbb{E}[\Delta x]}{\mathbb{E}[\tau]} = -\frac{0.271}{2.441} = -0.111, \tag{54}$$

and the volatility of idiosyncratic shocks equals

$$\sigma = \sqrt{\underbrace{\frac{\mathbb{E}[\Delta x^2]}{\mathbb{E}[\tau]}}_{0.08} + \underbrace{2\nu\hat{x}}_{0.002}} = 0.27. \tag{55}$$

Note that the main component that drives the volatility estimate is the second moment of capital gap changes, normalized by expected duration, whereas the drift term is negligible. The calibration of the volatility of innovations in the literature falls within a wide range, from 0.054 in Khan and Thomas (2008) to 0.121 in Winberry (2019) to 0.176 in Bachmann, Caballero and Engel (2013).[29] It is worth noting that these calibrations are done jointly with the fixed adjustment cost within a particular inaction model. In contrast, our volatility estimate is pinned down directly through our model–independent mapping between data and parameters.

---

concerns. We also present several robustness checks regarding the depreciation rates used in the calculation of capital stock via the PIM, the inaction threshold $\underline{i}$, the truncation for outliers, and adding Colombia as an additional country.

[29]To compare with the literature, we must do a few calculations. These papers use an AR(1) process for productivity in levels with volatility σ and autocorrelation ρ. If the process is calibrated at the annual frequency, then the annual standard deviation of productivity growth rate is $\sigma\sqrt{2/(1+\rho)}$; if the process is calibrated at the quarterly frequency, then the annual standard deviation of productivity growth rate is $\sigma\sqrt{\frac{\rho^4-1}{1-\rho^2} + 1 + \rho^2 + \rho^4 + \rho^6}$. Moreover, since we abstract from labor and our productivity is rescaled, we must adjust the previous numbers by a factor $1/(1-\alpha)$ in order to make them comparable to ours (we assume a labor share and span of control to deliver $\alpha = 0.58$). Lastly, for Bachmann, Caballero and Engel (2013), we only consider the idiosyncratic shocks (excluding the sectorial shocks). Recall that we only consider structures here.

Lastly, the observation formula for the reset capital gap in (22) implies that upon adjustment, capital gaps are reset on average 1.3% above the average capital gap:

$$\hat{x} = \underbrace{\frac{\mathbb{E}[\Delta x]}{2}}_{0.14} \underbrace{\left(1 - \mathbb{CV}[\tau]^2\right)}_{-0.23} + \underbrace{\frac{\mathbb{C}ov[\tau, \Delta x]}{\mathbb{E}[\tau]}}_{0.02} = 0.013 \tag{56}$$

As with the covariance, the reset state, together with the drift, provides highly useful information to distinguish families of inaction models, as we explain below.

Output from theory: Ergodic moments. From (24), average age $\mathbb{E}[a]$ is recovered using information about the average and the dispersion of adjustment times. Following our earlier discussion on renewal theory—i.e., larger stopping times are more representative in the sample—the heterogeneity in expected times increases the average age:

$$\mathbb{E}[a] = \underbrace{\mathbb{E}[\tau]}_{2.4} \underbrace{\frac{1 + \mathbb{CV}[\tau]^2}{2}}_{1.05} = 2.52. \tag{57}$$

Notice that $\mathbb{E}[a] \approx \mathbb{E}[\tau]$, and recalling an earlier discussion, this is consistent with a strong time-dependent component, where τ follows either a fixed date or an exponential distribution.

According to the observation formula (25), the steady-state dispersion of capital gaps $\mathbb{V}ar[x]$—a notion of misallocation—can be expressed in terms of capital gap changes and the reset point as follows:

$$\mathbb{V}ar[x] = \frac{\hat{x}^3 - \mathbb{E}[(\hat{x} - \Delta x)^3]}{3\mathbb{E}[\Delta x]} = 0.23, \tag{58}$$

where the cubic powers capture asymmetries in the distribution. Note that the ergodic variance is quantitatively very close to the average accumulated shocks, i.e., $\mathbb{V}ar[x] = \sigma^2 \mathbb{E}[a]$. This suggests a large pass-through from productivity shocks to capital misallocation, signaling inefficient capital adjustments.

Equation (26) implies that the covariance between age and adjustment size suggested by the data is positive:

$$\mathbb{C}ov[x, a] = \frac{1}{2\nu} \left(\underbrace{\frac{\mathbb{E}\left[\tau \left(\hat{x} - \Delta x\right)^2\right]}{\mathbb{E}[\tau]}}_{0.22} - \underbrace{\mathbb{V}ar[x]}_{0.23} - \underbrace{\sigma^2 \mathbb{E}[a]}_{0.20} \right) = 0.914. \tag{59}$$

This positive covariance between capital gaps and capital age means that the capital holdings of plants that have not adjusted in a long time (their capital is old) are *above* the gap of plants that have recently adjusted. In the next section, we discuss how this covariance, as well as other objects computed above, allows us to discern across models of inaction.

Let us recall the only two assumptions we have made to establish the link between data and ergodic moments and parameters: (i) capital gaps follow a Brownian motion, and (ii) the reset state $\hat{x}$ is constant across plants and time. We have not assumed a particular inaction model or a parametric restriction of the plants' state besides those imposed on the capital gaps x. Clearly, plants' investments may have

other drivers besides capital gaps, and we do not impose any structure on those.

4.4 Evidence in favor of random fixed adjustment costs

The structural connections between ergodic moments, parameters, and policies recovered from the data through our formulas strongly suggest that adjustment costs have both a time- and a state-dependent component, as well as a large asymmetry in the policy. Here, we summarize these pieces of evidence.

A fully state-dependent model can generate asymmetric policies that deliver positive covariance $\mathbb{C}ov[a, \Delta x]$ and the small and positive $\hat{x}$. However, such a model cannot generate the large observed passthrough $\mathbb{V}[x] = \sigma^2 \mathbb{E}[a]$. To see this, assume that such a model is able to generate the expected duration and the reset state; then it generates a variance of $\mathbb{V}[x] = 0.0020$, 100 times lower than in the data. In contrast, a fully time-dependent model can generate the large passthrough, but it faces two challenges: the lack of asymmetries needed to generate the positive covariance[30] (by definition, such models are symmetric) and the low reset state. If the driver of inaction is time-dependence, then the reset state should be equal to $\hat{x}^{\text{time dep}} = -\nu \mathbb{E}[a] = -0.111 * 2.554 = 0.279$, 20 times larger than the data.

For the previous reasons, we will focus on the random fixed adjustment cost to analyze aggregate dynamics.

4.5 CIR under random fixed adjustment costs

We start by assuming the Bernoulli fixed-cost model to analyze aggregate dynamics of the first moment of capital gaps. We consider an unanticipated permanent aggregate productivity shock that shifts horizontally the distribution of idiosyncratic productivity of all plants. If this model is true, in which both intensive and extensive margins are active, we obtain from the data the following values:

$$\text{CIR}_1(\delta)/\delta \approx \underbrace{\frac{\mathbb{V}ar[x]}{\sigma^2}}_{3.014} - \underbrace{\frac{\nu \, \mathbb{C}ov[a, x]}{\sigma^2}}_{1.340} = 4.354. \tag{60}$$

A natural question arises: Which parameters in the random fixed-cost model generate the data? To answer this, we use the following strategy. We explore three types of hazard rates. The first is a constant hazard generated by the baseline random fixed-cost model from the previous sections. We show that this model dramatically misses the data. Then we use the generalized hazard model to entertain two simple deviations that can be parametrized easily: (i) a U–shaped hazard, as proposed by Caballero and Engel (1999) and followed by most of the literature, and (ii) a piecewise linear hazard. The advantage of these formulations is that quantitatively, the CIR computed under these two alternatives is well approximated by the variance and covariance expression in (37); thus we keep the focus on these two moments. In each of these cases, we fix the values for $(\sigma, \nu, \hat{x}, \mathbb{E}[\tau])$ obtained in the data and then search for a parametrization that best fits the variance and covariance. We conclude that a piecewise linear hazard provides the best fit for the data.

---

[30]In fully time-dependent models the covariance equals $\mathbb{C}ov[a, \Delta x] = \frac{\nu}{\mathbb{E}[\tau]}\left(\frac{\mathbb{E}[\tau^2]}{2} + \frac{\mathbb{E}[\tau^3]}{3}\right) < \frac{\nu}{\mathbb{E}[\tau]}\frac{\mathbb{E}[\tau^2]}{2} = -0.22$, which is extremely low compared to its value in the data

Bernoulli fixed cost (constant hazard). We search values for $(\lambda, \underline{x}, \overline{x})$. Fix the hazard's level λ. Find the inaction region's bounds $\underline{x}(\lambda), \overline{x}(\lambda)$ to match the average inaction duration $\mathbb{E}[\tau]$ and the reset state $\hat{x}$. Compute the implied variance $\mathbb{V}ar[x]$ and covariance $\mathbb{C}ov[x, a]$ and repeat for other values of λ. Panels A and C in Figure IV plot the two key moments and the CIR$_1$ against λ. It also shows the values in the data (dashed lines). The variance and covariance are increasing, and the combination that yields the best match is $(\lambda, \underline{x}, \overline{x}) = (0.217, -0.55, \infty)$. The calibration pushes for a very large and positive upper border of inaction $\overline{x}$, which in turn makes the lower border more negative to fit the reset state and the frequency of free adjustment λ larger. At some point the upper border becomes irrelevant and λ and $\underline{x}$ converge to the values above. The implied values are far from the data: a variance of $\mathbb{V}ar[x] = 0.02$ and a covariance of $\mathbb{C}ov[x, a] = 0.3$, implying a very low propagation (only 14% of the propagation in the data). As the next step, we use the general hazard to explore simple departures.

U-shaped hazard. We have shown that a random fixed-cost model with constant hazard is not able to match the two key moments in the data and thus misses the CIR. We entertain a first deviation in which the fixed-cost is distributed uniformly and generates a quadratic hazard of the form

$$\Lambda(x) = c(x - \hat{x})^2, \qquad c > 0. \tag{61}$$

Given the parameters for the stochastic process, we search values for $(c, \underline{x}, \overline{x})$ so that this model gets as close as possible to the data. Fix a quadratic coefficient c. Find the inaction region's bounds $\underline{x}(c), \overline{x}(c)$ to match the average inaction duration $\mathbb{E}[\tau]$ and the reset state $\hat{x}$. Then compute the implied capital gap variance $\mathbb{V}ar[x]$ and covariance with age $\mathbb{C}ov[x, a]$. Repeat for other values of c. Panels B and E plot the two key moments and the CIR for different values of the quadratic coefficient c. The combination that yields the best match is $(c, \underline{x}, \overline{x}) = (3.05, -0.44, 1.05)$, but again, the implied values are still far from the data: a variance of $\mathbb{V}ar[x] = 0.04$ and a covariance of $\mathbb{C}ov[x, a] = 0.03$. Notice that the CIR is remarkably well approximated by the variance and covariance formula. By missing the moments in the data, this model also implies very little propagation of aggregate shocks (again, only 14%).

Piecewise linear hazard. Finally, we analyze a piecewise linear hazard of the form

$$\Lambda(x) = \max\{b_0 + b_1(x - \hat{x}), 0\}, \qquad b_0 > 0, \ b_1 < 0. \tag{62}$$

This hazard incorporates strong asymmetries into the adjustment cost structure, as $\Lambda(x) = 0$ for all sufficiently large gaps: $x > \hat{x} - b_0/b_1$. This hazard resembles a 'hockey stick'. In other words, if the capital gap is too positive, there is no disinvestment (downward irreversibility). Given the parameters for the stochastic process, we search values for $(b_0, b_1, \underline{x})$. Inspired by the estimation of the constant hazard, the inaction region's upper bound is set at $\overline{x} = \infty$. Fix a hazard slope b_1. Find the intercept $b_0(b_1)$ and the inaction region's lower bound $\underline{x}(b_1)$ to match the average inaction duration $\mathbb{E}[\tau]$ and the reset state $\hat{x}$. Compute the variance $\mathbb{V}ar[x]$ and covariance $\mathbb{C}ov[x, a]$ and repeat for other values of b_1. Panels C and F plot the two key moments and the CIR for different values of the hazard's slope b_1. The combination that yields the best match is $(|b_1|, b_0, \underline{x}) = (0.74, 0.37, -1.7)$, and the implied values for the

variance $\mathbb{V}ar[x] = 0.17$ and the covariance $\mathbb{C}ov[x, a] = 0.6$ are much closer to the data. As in the previous case, the CIR is approximated perfectly by the variance and covariance formula. This model is able to explain 75% of the propagation in the data.

Figure IV – Key moments and CIR in data and models with random costs (three different hazards)
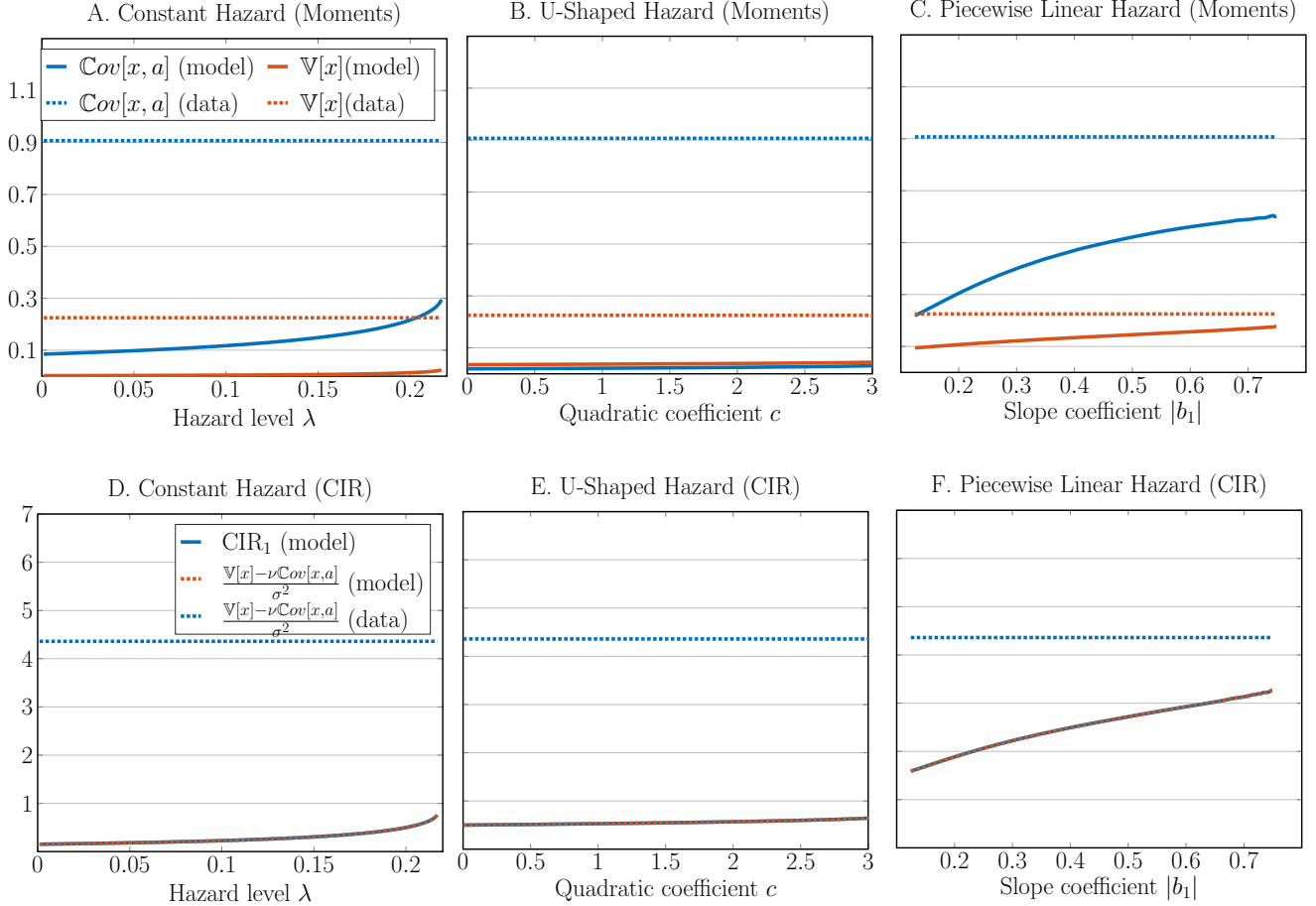
Notes: Panels A, B, and C plot the steady-state variance of capital gaps and the covariance between capital gap and age for three random fixed-cost models: Bernoulli (constant hazard), U-shaped hazard, and piecewise linear hazards. Each line describes the curve level across one parameter such that the models satisfy $(\nu, \sigma, \hat{x}, \mathbb{E}[\tau]) = (-0.111, 0.27, 0.013, , 2.44)$. Panels D, E, and F plot the CIR_1 in each model, together the implied CIRs computed as $\frac{\mathbb{V}[x] - \nu \mathbb{C}ov[x, a]}{\sigma^2}$ with the moments computed in the data and within each model.

On the nature of adjustment costs. We conclude that in order to explain aggregate investment dynamics, a lumpy investment model must match two cross-sectional moments: the variance and the covariance. In order to get close to reproducing these two moments in the data, the standard lumpy model must be augmented to feature two elements: downward irreversibility for sufficiently large investments and sufficient randomness in the fixed cost.

4.6 Robustness

We conduct a series of robustness checks for the empirical analysis. Overall, values for the key statistics of interest, $\mathbb{V}ar[x]$ and $\mathbb{C}ov[x, a]$, are very similar across all the following alternatives. A first concern that arises is how different layers of heterogeneity may affect the computation and interpretation of cross-sectional statistics. We recompute the statistics controlling for heterogeneity in terms of capital type (structures, vehicles, machinery, and equipment), plant size (number of workers), and industrial sectors; we also compare unweighted vs. weighted statistics. Second, we address the potential presence of time trends and differences along the business cycle phase. Third, we explore alternative methodologies to construct capital gap adjustments. Finally, we repeat the entire analysis for Colombia.

Heterogeneity by capital category. We repeat the analysis for each capital category separately: structures, machinery, equipment, and vehicles. Table X in the Data Appendix reports cross-sectional investment statistics by capital type, and Figure XI plots the histograms of capital gaps. Table XI reports data inputs and theory outputs. Across all categories the covariance is positive, but it is almost zero for vehicles. This is not surprising, as vehicles are not expected to have strong irreversibility given a well-developed secondary market.

Heterogeneity by sector. We consider eight 2-digit sectors within manufacturing: (1) Food and beverages; (2) Textiles, clothing and leather; (3) Wood and furniture; (4) Paper and printing; (5) Chemistry, petroleum, rubber and plastic; (6) Manufacture of non-metallic mineral products; (7) Basic metal; and (8) Metal products, machinery and equipment. Table XII in the Data Appendix reports cross-sectional statistics by sectors and Tables XIII and XIV shows data inputs and theory outputs for structures and total capital, respectively. We find that besides textiles and chemicals, all other sectors present similar investment patterns. When we recompute the formulas excluding textiles and chemicals, we find that neither the inputs from the data nor the outputs from the theory differ from the numbers that include these sectors. We conclude that heterogeneity across sectors should not be a concern once these two sectors are excluded.

Heterogeneity by plant size. We consider different plant sizes in terms of their number of workers. Table XV in the Data Appendix reports cross-sectional statistics by quartiles of the average number of workers during the sample period: small plants (0–25%, S), medium plants (25–50%, M), large plants (50–75%, L), and very large plants (75–100%, XL). Then quantiles are computed using a plant-level fixed effect and a time trend. Table XVI computes the formulas by plant size. Inputs from the data are different for small and large plants. What is particularly striking is that the parameters of the stochastic process we recovered are identical across sizes: plants are hit with the same magnitude of shocks and the same drift. Variance and covariance are also very similar (positive and large) across plant size, even if the level of lumpiness differs dramatically across them.

Time trends and business cycle. To address the potential presence of time trends that may contaminate the statistics, we recompute the formulas for three 10–year subperiods (1979–1989, 1990–2000,

2001–2011). To assess whether the parameters and moments change along the business cycle, we compute the formulas conditioning on the business cycle phase, where recession periods are identified following the OECD index of economic indicators. Results are in Table XVII and XVIII. While we find the presence of some trends, the moments of interest remain large and positive.

Weighted statistics and unbalanced panel. We consider weighted statistics by plants' total capital and also an unbalanced panel. We repeat all of the configurations above. The results in Table XIX show the robustness to these alternative specifications.

Alternative construction of capital gaps. We consider alternative depreciation rates to construct the capital stock through the perpetual inventory method $\{0.5\delta, \delta, 1.5\delta\}$, where δ varies by capital type, as in the baseline. Results are in Table XX. Additionally, we consider different cut values to define inaction in equation (52): $\underline{i} \in \{0, 0.01, 0.02\}$. We also explore different truncation levels for outliers, from 2% (1st and 99th percentiles) to 4% (2nd and 98th). Results are in Table XXI.

Additional country: Colombia. We repeat all analyses for Colombia, using the Annual Manufacturers Survey (*Encuesta Anual Manufacturera*) for the period 1995–2016. These data has been used by Eslava, Haltiwanger, Kugler and Kugler (2004, 2013) to study the effect of structural reforms and trade liberalization on aggregate productivity. Results are very similar across the two countries (see Data Appendix E).

5 Generalization and extensions

In the previous sections we specified parametric restrictions to the inaction model and to the firms' state space to get expressions for the CIR. Nevertheless, it is possible to extend our theory to accommodate richer models. For instance, models with both observation and fixed costs, as in Álvarez, Lippi and Paciello (2011), can be represented by adding new state variables. In this section, we generalize our results to consider more general stopping rules and states, explaining the assumptions on policies and processes that are necessary in order to apply our tools. We extend the analysis in four directions, to consider: (i) fixed heterogeneity, (ii) a mean-reverting process for the state, (iii) transitions for arbitrary smooth functions of the state, and (iv) aggregate second-moment shocks.

5.1 Generalization

Notation. We denote conditional distributions as $Z|Y$, conditional expectations with initial condition z as $\mathbb{E}^z[Z]$, and the minimum between two stopping times as $t \wedge s \equiv \min\{t, s\}$.

General setup. Let $(\Omega, P, \mathcal{F})$ be a probability space equipped with a filtration $\mathcal{F} = (\mathcal{F}_t; t \geq 0)$. We consider an economy populated by a continuum of agents indexed with $\omega \in \Omega$, where agent ω's information set at time t is the filtration $\mathcal{F}_t$. Each agent's uncontrolled state is given by $\tilde{S}_t(\omega) = [\tilde{x}_t(\omega), S_t^{-x}(\omega)] \in \mathbb{R}^{1+K_{-x}}$. The state is split between a *main state* $\tilde{x}$ and a set of *complementary states* $\tilde{S}_t^{-x}$. The main

state follows a Brownian motion with drift $d\tilde{x}_t(\omega) = \nu dt + \sigma dW_t(\omega)$. The agent's policies consist of a sequence of adjustment dates $\{T_k\}_{k=1}^{\infty}$ and adjustment sizes $\{\Delta S_{\tau_k}\}_{k=1}^{\infty}$, measurable with respect to $\mathcal{F}_t$. Given these policies $\{T_k(\omega), \Delta S_{\tau_k}(\omega)\}_{k=1}^{\infty}$, the controlled state $S_t(\omega)$ evolves as the sum of the uncontrolled state plus the adjustments: $S_t(\omega) = \tilde{S}_t(\omega) + \sum_{T_k(\omega) \leq t} \Delta S_{\tau_k}(\omega)$.

The first premise for our theory is a recursive representation of the conditional CIR, both between and within stopping dates. This requires that $S_t(\omega)$ be a sufficient statistic for the conditional CIR, which in turn requires that the policy be history independent. Formally, this means that for all $t(\omega) \leq T_{i+1}(\omega)$ we have that

$$\mathbb{E}\left[\int_{T_i \wedge t(\omega)}^{T_{i+1}} f(x_t) dt \Big| \mathcal{F}_{T_i \wedge t(\omega)}\right] = \mathbb{E}\left[\int_0^{T_{i+1} - (T_i \wedge t(\omega))} f(x_t) dt \Big| S_0 = S_{T_i \wedge t(\omega)}\right] = v^f(S_{T_i \wedge t(\omega)}). \tag{63}$$

Since the main state follows a Brownian motion, the burden of this requirement falls completely on the complementary state and the policy. Assumptions 1 and 2 formalize these requirements.

Assumption 1 (Markovian complementary state). *The complementary state $\tilde{S}_t^{-x}$ follows a strong Markov process:*

$$\tilde{S}_{(t \wedge T_k)+h}^{-x} | \mathcal{F}_{t \wedge T_k} = \tilde{S}_h^{-x} | \{\tilde{S}_0^{-x} = \tilde{S}_{(t \wedge T_k)}\}, \qquad \forall k. \tag{64}$$

To understand this assumption, consider a history ω such that $t < T_k(\omega)$. In this case, the complementary state's law of motion depends only on its current value; thus it is independent of its own history. Additionally, the complementary state is an homogeneous process, since its law of motion at date t is equivalent to its law of motion at zero, given an initial condition. In the complementary case $t \geq T_k(\omega)$, these properties continue to hold, and thus the stopping policy does not reveal new information about the complementary state's law of motion.

Assumption 2 (Markovian policies). *Policies satisfy the following conditions:*

$$T_{k+1} | \mathcal{F}_{T_k+h} = T_1 | \{S_0 = S_{T_k+h}\} \quad \text{for all} \quad h \in [0, T_{k+1} - T_k]. \tag{65}$$

A second premise in our theory is that we can characterize the CIR with the *first* stopping time of every agent. This means that, upon taking action, agents fully adjust to include any deviations from their steady-state behavior and come back to the steady-state process. This would imply that S_{τ_k} is *iid* across time and independent of the history previous to the adjustment. The challenge with stochastic *iid* resets is that is makes it more difficult to identify the parameters of the stochastic process, e.g., differentiating the fundamental volatility σ from the volatility arising from a random reset state. Therefore, in order for the reset state to be sufficiently informative, we ask that it be a constant $S_{T_i} = \hat{S}$.

Assumption 3 (Constant reset state). *The reset state is constant: $S_{T_k} = \hat{S}$ for all k.*

Finally, in order to apply the Optional Sampling Theorem, we require several stopping processes to be well-defined (finite moments at the stopping-time).[31]

---

[31]See Online Appendix A for a formal definition of a well-defined stopping process.

Assumption 4 (**Well-defined stopping processes**). *The processes* $\left(\left\{ \int_0^t s^j x_s^m \, dB_s \right\}_t, \tau \right)$ *for all* m *and* $j = 0, 1$ *are well-defined stopping processes.*

It is straightforward to check that the previous assumptions hold in the investment example developed in Section 2.[32] Assumptions 1 to 4 are enough to demonstrate Propositions 1 and 2. Proposition 3 cannot be directly extended, since it depends on having the same basis underlying the CIR and the ergodic moments. Under the condition that the stopping policy is independent of the main state, we can obtain the CIR characterization in Proposition 4. In order to extend Proposition 5, we require one additional assumption. There must exist an equivalent representation of the extensive margin as a function exclusively of the main state x. For this, we require that there exist a stopping policy τ^* that only depends on the main state x and can fully describe the extensive margin by itself. For instance, a stopping policy given by a Poisson counter with hazard $\Lambda(x)dt$ satisfies this requirement.

Assumption 5 (**Hazard**). *There exists a stopping policy* τ^* *s.t.*

$$\mathbb{E}^{\hat{S}} \left[\int_0^\tau \left(\frac{\partial \mathbb{E}^S \left[x_\tau^{m+1}/m + 1 \right]}{\partial x} - \mathbb{E}^S \left[x_\tau^m \right] \right) dt \right] = \mathbb{E}^{\hat{x}} \left[\int_0^\tau \left(\frac{\partial \mathbb{E}^x \left[x_{\tau^*}^{m+1}/m + 1 \right]}{\partial x} - \mathbb{E}^x \left[x_{\tau^*}^m \right] \right) dt \right]. \quad (66)$$

With the formal requirements needed to apply our theory, we proceed to develop four extensions.

5.2 Extensions

Extension I: Fixed heterogeneity. The first extension considers fixed heterogeneity across different types of firms, e.g., sectors. Let $n = 1, \ldots, N$ denote the sectors, (ν_n, σ_n) the sectoral drift and volatility, and γ_n the output weight of sector n. Then, up to the first order, the CIR_m of the economy is a weighted average of the sector-specific $\text{CIR}_{m,n}$:[33]

$$\text{CIR}_m(\delta) = \sum_{n=1}^N \gamma_n \text{CIR}_{m,n}(\delta)). \quad (67)$$

The only important consideration for this derivation is that capital gaps refer to the capital-productivity ratio relative to the sectoral average, and *not* the economy's average. Note that, within each sector, (ν_n, σ_n) is a constant vector. If this is not the case, the theory can still recover cross-sectional moments of these parameters. For example, under a scenario with heterogeneous volatility and zero drift, Baley and Blanco (2019) show how to recover average volatility across agents: $\mathbb{E}[\sigma_\omega^2] = \mathbb{E}[\Delta x^2]/\mathbb{E}[\tau]$.

Extension II: Mean reversion. Mean reverting processes are widely used in many applications due to their empirical relevance. Assume the uncontrolled state follows an Ornstein–Uhlenbeck process with mean-reversion parameter ρ, i.e., $d\tilde{x}_t = \rho \tilde{x}_t dt + \sigma dW_t$. We show that the theoretical tools that map

---

[32]For Assumption 1, the complementary state is given by the arrival of free adjustment opportunities N_t, which is assumed to be a Poisson counter process and thus a strong Markov process. The requirements in Assumptions 2 and 3 are also satisfied. We show that the reset capital gap is constant; and since the stopping policy is an inaction set with respect to the controlled state, the stopping policy is history independent within and between adjustments.

[33]See Online Appendix B.3.

aggregate dynamics to steady-state moments to micro data still apply, but the precise formulas change. For instance, the reset state and structural parameters are recovered through the following system:[34]

$$\hat{x} = \frac{\mathbb{E}[e^{-\rho\tau}\Delta x]}{\mathbb{E}[e^{-\rho\tau}] - 1}, \tag{68}$$

$$\sigma^2/\rho = 2\frac{\hat{x}^2 - \mathbb{E}\left[e^{-2\rho\tau}(\hat{x} - \Delta x)^2\right]}{\mathbb{E}\left[e^{-2\rho\tau}\right] - 1}, \tag{69}$$

$$erf\left(\frac{\hat{x}}{\sqrt{\sigma^2/\rho}}\right) = \mathbb{E}\left[erf\left(\frac{\hat{x} - \Delta x}{\sqrt{\sigma^2/\rho}}\right)\right]. \tag{70}$$

where $erf(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the Gauss error function. Note that in the limiting case $\rho \to 0$, and using the approximation $e^{-\rho\tau} \approx 1 - \rho\tau$, expressions (68) and (69) converge to expressions in (22) and (23) with zero drift $\nu = 0$. These equations are useful to test for the persistence of productivity shocks directly from investment data.

Extensions III: Transitions of arbitrary functions of the state. The third extension studies transitions of any arbitrary smooth functions of the state $f(x)$. This extension can be applied to study welfare, as the welfare criteria can be generally written in this form. Additionally, it can be used to correct for biases arising from applying logs and aggregating.[35] Using a Taylor approximation around zero, we write the CIR of the $f(x)$ function in terms of a sequence of CIRs, weighted by the Taylor factors:

$$\text{CIR}(f(x), \delta) = \int_0^\infty \mathbb{E}_t[f(x)] - \mathbb{E}[f(x)]dt = \sum_{j=1}^\infty \frac{df^j(0)}{dx^j}\frac{\text{CIR}_j(\delta)}{j!}. \tag{71}$$

Extension IV: Aggregate second-moment shocks. Since the seminal work on aggregate uncertainty shocks by Bloom (2009), there has been great interest in the macroeconomic consequences of second-moment shocks along the cycle. For simplicity, we consider a perturbation that can be expressed via a single parameter δ such that the initial distribution $F_0(x) = F(x\delta)$ represents a mean-preserving spread ($\delta < 1$) or contraction ($\delta > 1$) of the ergodic distribution. To simplify the exposition, we set $\nu = 0$ and focus on the dynamics of the first moment[36]

$$\text{CIR}_1(\delta)/\delta = \Gamma + \Psi + o(\delta) \tag{72}$$

$$\Gamma = 2\mathbb{C}ov[x, a] \tag{73}$$

$$\Psi = \sum_{j=0}^\infty \psi_{1,j}\mathbb{E}[x^{j+1}] \tag{74}$$

As with first-moment shocks, the $\mathbb{C}ov[x, a]$ is also a key statistic to study aggregate dynamics following second-moments shocks. Intuitively, an aggregate shock to the second moment of the distribution can

---

[34]See Online Appendix B.4.

[35]Campbell and Fisher (2000) show that working with the average of the log (instead the log of the average) generates a bias in aggregate dynamics as agents with different sizes are treated equality in the log-scale. Computing the CIR for $f(x) = e^x$ corrects this bias.

[36]See Online Appendix B.5.

affect the dynamics of the first moment under large asymmetries in the policy, which are captured by the covariance term. This term could be positive or negative, and it is disciplined by the data. The interaction between first- and second-moment shocks, as studied by Vavra (2014), Castelnuovo and Pellegrino (2018), and Baley and Blanco (2019), can be easily accommodated by combining our results.

6 Conclusion

We provide structural relationships in lumpy models between the CIR—a summary measure of persistence for aggregate dynamics—steady-state cross-sectional moments, and micro data on adjustments, and apply them to lumpy investment. We conclude by discussing three promising avenues for future research that may overcome some of the limitations in our work.

First, we focus on a one-dimensional state. Extending the theory to multidimensional states would allow us to analyze the interaction between heterogeneity and lumpiness in multi-plant firms (Kehrig and Vincent, 2019), portfolio choice with multiple assets (Kaplan, Moll and Violante, 2018), misallocation with several production inputs (Hawkins, Michaels and Oh, 2015), or the interaction of lumpy behavior across different choices, such as investment and price-setting (Sveen and Weinke, 2007).

Second, we assume full adjustment upon action. Extending the theory to accommodate partial adjustments would allow us to study the interaction of lumpiness with convex adjustment costs or imperfect information. This line of work would continue the recent contributions by Baley and Blanco (2019).

Third, we characterize the CIR but not the complete IRF, and we only consider small perturbations around steady-state. Extending the theory to characterize the full IRF and more general perturbations, as in contemporaneous work by Álvarez and Lippi (2019) who make significant progress in this front using eigenvalue-eigenfunction representations, would allow us to better understand the role of lumpiness for aggregate dynamics.

References

ÁLVAREZ, F., BERAJA, M., GONZALEZ-ROZADA, M. and NEUMEYER, P. A. (2018). From hyperinflation to stable prices: Argentina's evidence on menu cost models. *The Quarterly Journal of Economics*, **134** (1), 451–505.

—, LE BIHAN, H. and LIPPI, F. (2016). The real effects of monetary shocks in sticky price models: a sufficient statistic approach. *American Economic Review*, **106** (10), 2817–51.

— and LIPPI, F. (2014). Price setting with menu cost for multiproduct firms. *Econometrica*, **82** (1), 89–135.

— and LIPPI, F. (2019). The analytic theory of a monetary shock. *EIEF Working Paper*.

—, LIPPI, F. and PACIELLO, L. (2011). Optimal price setting with observation and menu costs. *Quarterly Journal of Economics*.

—, — and PACIELLO, L. (2016). Monetary shocks in models with inattentive producers. *The Review of Economic Studies*, **83** (2), 421–459.

ALVAREZ, F., LIPPI, F. and PASSADORE, J. (2017). Are state-and time-dependent models really different? *NBER Macroeconomics Annual*, **31** (1), 379–457.

ASKER, J., COLLARD-WEXLER, A. and DE LOECKER, J. (2014). Dynamic inputs and resource (mis) allocation. *Journal of Political Economy*, **122** (5), 1013–1063.

BACHMANN, R. and BAYER, C. (2014). Investment dispersion and the business cycle. *American Economic Review*, **104** (4), 1392–1416.

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

BALEY, I. and BLANCO, A. (2019). Firm uncertainty cycles and the propagation of nominal shocks. *AEJ: Macro*, **11** (1), 276–337.

BAYER, C. (2009). A comment on the economics of labor adjustment: Mind the gap: Evidence from a monte carlo experiment. *American Economic Review*, **99** (5), 2258–66.

BERGER, D. and VAVRA, J. (2019). Shocks vs responsiveness: What drives time-varying dispersion. *Journal of Political Economy*, **Forthcoming**.

BERRY, S., LEVINSOHN, J. and PAKES, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pp. 841–890.

BERTOLA, G. and CABALLERO, R. J. (1994). Irreversibility and aggregate investment. *The Review of Economic Studies*, **61** (2), 223–246.

BLANCO, A. (2019). Optimal inflation target in an economy with menu costs and an occasionally binding zero lower bound. *Working Paper*.

— and CRAVINO, J. (2018). Price rigididity and the relative ppp. *NBER Working Paper*.

BLOOM, N. (2009). The impact of uncertainty shocks. *Econometrica*, **77** (3), 623–685.

—, GUVENEN, F. and SALGADO, S. (2016). Skewed business cycles. In *2016 Meeting Papers*, 1621, Society for Economic Dynamics.

BOROVIČKA, J., HANSEN, L. P. and SCHEINKMAN, J. A. (2014). Shock elasticities and impulse responses. *Mathematics and Financial Economics*, **8** (4), 333–354.

BURSTEIN, A. and HELLWIG, C. (2008). Welfare costs of inflation in a menu cost model. *The American Economic Review*, pp. 438–443.

CABALLERO, R. J. and ENGEL, E. (1991). Dynamic (s,s) economies. *Econometrica*, **59** (6), 1659–1686.

— and — (1999). Explaining investment dynamics in u.s. manufacturing: A generalized (s, s) approach. *Econometrica*, **67** (4), 783–826.

—, — and HALTIWANGER, J. (1995). Plant-level adjustment and aggregate investment dynamics. *Brookings papers on economic activity*, pp. 1–54.

—, — and — (1997). Aggregate employment dynamics: Building from microeconomic evidence. *The American Economic Review*, **87** (1), 115–137.

— and ENGEL, E. M. (2007). Price stickiness in ss models: New interpretations of old results. *Journal of monetary economics*, **54**, 100–121.

CALVO, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, **12** (3), 383–398.

CAMPBELL, J. R. and EDEN, B. (2014). Rigid prices: Evidence from us scanner data. *International Economic Review*, **55** (2), 423–442.

— and FISHER, J. D. M. (2000). Aggregate employment fluctuations with microeconomic asymmetries. *The American Economic Review*, **90** (5), 1323–1345.

CAPLIN, A. and LEAHY, J. (1991). State-Dependent Pricing and the Dynamics of Money and Output. *The Quarterly Journal of Economics*, **106** (3), 683–708.

— and — (1997). Aggregation and optimization with state-dependent pricing. *Econometrica: Journal of the Econometric Society*, pp. 601–625.

CAPLIN, A. S. and SPULBER, D. F. (1987). Menu costs and the neutrality of money. *The Quarterly Journal of Economics*, **113**, 287–303.

CARVALHO, C. and NECHIO, F. (2011). Aggregation and the ppp puzzle in a sticky-price model. *American Economic Review*, **101** (6), 2391–2424.

— and SCHWARTZMAN, F. (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics*, **76**, 141–156.

CASTELNUOVO, E. and PELLEGRINO, G. (2018). Uncertainty-dependent effects of monetary policy shocks: A new keynesian interpretation. *Journal of Economic Dynamics and Control*, **Forthcoming**.

COOPER, R. and WILLIS, J. L. (2004). A comment on the economics of labor adjustment: Mind the gap. *The American Economic Review*, **94** (4), 1223–1237.

COOPER, R. W. and HALTIWANGER, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, **73** (3), 611–633.

DAVID, J. M. and VENKATESWARAN, V. (2019). The sources of capital misallocation. *American Economic Review*, **Forthcoming**.

DE LOECKER, J. and EECKHOUT, J. (2017). The rise of market power and the macroeconomic implications, uCL mimeo.

— and WARZYNSKI, F. (2012). Markups and firm-level export status. *American economic review*, **102** (6), 2437–71.

DIXIT, A. K. and PINDYCK, R. (1994). *Investment under uncertainty*. Princeton university press.

EHOUARNE, C., KUEHN, L.-A. and SCHREINDORFER, D. (2016). *Misallocation cycles*. Tech. rep., Working Paper.

ELSBY, M., RATNER, D. and MICHAELS, R. (2019). The aggregate effects of labor market frictions. *Quantitative Economics*, **Forthcoming**.

ELSBY, M. W. and MICHAELS, R. (2019). Fixed adjustment costs and aggregate fluctuations. *Journal of Monetary Economics*, **101**, 128–147.

ESLAVA, M., HALTIWANGER, J., KUGLER, A. and KUGLER, M. (2004). The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from colombia. *Journal of Development Economics*, **75** (2), 333 – 371, 15th Inter American Seminar on Economics.

—, —, — and — (2013). Trade and market selection: Evidence from manufacturing plants in colombia. *Review of Economic Dynamics*, **16** (1), 135 – 158, special issue: Misallocation and Productivity.

FEENSTRA, R. C., INKLAAR, R. and TIMMER, M. P. (2015). The next generation of the penn world table. *American Economic Review*, **105** (10), 3150–82.

GAGNON, E. (2009). Pricing setting during low and high inflation: Evidence from mexico. *Quarterly Journal of Economics*, p. 126.

GOLOSOV, M. and LUCAS, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, pp. 171–199.

GOURIO, F. and KASHYAP, A. K. (2007). Investment spikes: New facts and a general equilibrium exploration. *Journal of Monetary Economics*, **54**, 1–22.

HAMERMESH, D. S. (1989). Labor demand and the structure of adjustment costs. *The American Economic Review*, **79** (4), 674–689.

HAWKINS, W., MICHAELS, R. and OH, J. (2015). *The joint dynamics oF capital and employment at the plant level*. Tech. rep., Mimeo, Yeshiva University.

HENRÍQUEZ, C. G. (2008). *Stock de Capital en Chile (1985-2005): Metodologa y Resultados*. Economic Statistics Series 63, Central Bank of Chile.

HSIEH, C.-T. and KLENOW, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, **124** (4), 1403–1448.

KAPLAN, G., MOLL, B. and VIOLANTE, G. L. (2018). Monetary policy according to hank. *American Economic Review*, **108** (3), 697–743.

KARADI, P. and REIFF, A. (2019). Menu costs, aggregate fluctuations, and large shocks. *AEJ: Macro*, **Forthcoming**.

KEHRIG, M. and VINCENT, N. (2019). Good dispersion, bad dispersion. *Duke Working Paper*.

KHAN, A. and THOMAS, J. K. (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, **76** (2), 395–436.

LIU, L. (1993). Entry-exit, learning, and productivity change evidence from chile. *Journal of Development Economics*, **42** (2), 217 – 242.

MANKIW, N. G. (2014). *Principles of macroeconomics*. Cengage Learning.

MIDRIGAN, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica*, **79** (4), 1139–1180.

NAKAMURA, E. and STEINSSON, J. (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly Journal of Economics*, **125** (3), 961–1013.

OBERFIELD, E. (2013). Productivity and misallocation during a crisis: Evidence from the chilean crisis of 1982. *Review of Economic Dynamics*, **16** (1), 100–119.

OKSENDAL, B. (2007). *Stochastic Differential Equations*. Springer, 6th edn.

POMBO, C. (1999). Productividad industrial en colombia: una aplicación de números índices. *Revista de economía del Rosario*, **2** (1), 107–139.

REIS, R. (2006). Inattentive Producers. *The Review of Economic Studies*, **73** (3), 793–821.

STOKEY, N. (2009). *The Economics of Inaction*. Princeton University Press.

SVEEN, T. and WEINKE, L. (2007). Lumpy investment, sticky prices, and the monetary transmission mechanism. *Journal of Monetary Economics*, **54**, 23–36.

TAYLOR, J. B. (1980). Aggregate dynamics and staggered contracts. *The Journal of Political Economy*, pp. 1–23.

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

TYBOUT, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic Literature*, **38** (1), 11–44.

VAVRA, J. (2014). Inflation dynamics and time-varying volatility: New evidence and an ss interpretation. *The Quarterly Journal of Economics*, **129** (1), 215–258.

VERACIERTO, M. L. (2002). Plant-level irreversible investment and equilibrium business cycles. *American Economic Review*, **92** (1), 181–197.

WINBERRY, T. (2019). Lumpy investment, business cycle, and stimulus policy. *American Economic Review*, **Forthcoming**.

WOODFORD, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics*, **56**, S100–S124.

ZWICK, E. and MAHON, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, **107** (1), 217–48.

A Appendix: Preliminaries

Assumptions. Let $S_t = [x_t, a_t]$ be the state. Here we enumerate three assumptions needed for Propositions 1–4:

1. $\{S_t\}_{t \in [0,\infty)}$ is an strong Markov process with first element x_t.

2. Homogeneous resets: $S_{\tau_i} = \hat{S}$.

3. τ_i is a stopping time with respect to the filtration generated by $\{S_t\}$.

4. $\left(\left\{ \int_0^t x_s^m s^n \, dW_s \right\}_t, \tau \right)$ are well-defined stopping processes for any m and $n = 0, 1$.

Auxiliary Theorem 1. *[**Optional Sampling Theorem (OST)**] Let Z be a (sub) martingale on the filtered space $(\Omega, \mathcal{P}, \mathcal{F})$ and τ a stopping time. If $(\{Z_t\}_t, \tau)$ is a well-defined stopping process, then*

$$\mathbb{E}[Z_\tau](\geq) = \mathbb{E}[Z_0] \tag{A.1}$$

Proof. See Theorem 4.4 in Stokey (2009). □

Auxiliary Theorem 2. *[**Ergodic distribution and occupancy measure**] Let S be a strong Markov process and $g : S \to \mathbb{R}$ a function of S. Denote with F the ergodic distribution of S and with R the renewal distribution, i.e. the distribution conditional on adjustment. With homogeneous resets, $\Pr[S = \hat{S}] = 1$ under the renewal distribution R. Assume $\int g(S) dF(S) = \lim_{T \to \infty} \frac{\int_0^T g(S_t) dt}{T}$ for all initial conditions S_0. Then the following relationship holds:*

$$\int g(S) dF(S) = \frac{\mathbb{E}\left[\int_0^\tau g(S_t) dt \right]}{\mathbb{E}[\tau]}. \tag{A.2}$$

Proof. See proof in Online Appendix A . □

B Appendix: Proofs

Proof of Proposition 1. We continue with the notation $S_t = [x_t, a_t]$ and Assumptions 1 to 4 above. Additionally, we assume that the first element of S_t follows a Brownian motion with non-zero drift. For the zero-drift case, see Online Appendix B.2.

- **Average adjustment size.** From the law of motion $x_t = \hat{x} + \nu t + \sigma W_t$, we find the following equalities: $\sigma W_\tau = -\nu \tau + x_\tau - \hat{x} = -\nu \tau - \Delta x$. Taking expectations on both sides, we have $\sigma \mathbb{E}[W_\tau] = -\nu \mathbb{E}[\tau] - \mathbb{E}[\Delta x]$. Since W_τ is a martingale, $\mathbb{E}[W_\tau] = W_0 = 0$ by the OST. Therefore, $\nu = -\frac{\mathbb{E}[\Delta x]}{\mathbb{E}[\tau]}$ as well.

- **Fundamental volatility:** To characterize σ, define $Y_t = x_t - \nu t$ with initial condition $Y_0 = \hat{x}$. With similar steps as before, we have that

$$\sigma^2 = \frac{\mathbb{E}\left[\Delta Y_\tau^2\right]}{\mathbb{E}[\tau]} = \frac{\mathbb{E}\left[(Y_\tau - Y_0)^2\right]}{\mathbb{E}[\tau]} = \frac{\mathbb{E}\left[(x_\tau - \nu\tau - \hat{x})^2\right]}{\mathbb{E}[\tau]} = \frac{\mathbb{E}\left[(\nu\tau + \Delta x)^2\right]}{\mathbb{E}[\tau]} \tag{B.3}$$

  or equivalently

$$\sigma^2 = \frac{\mathbb{E}[\Delta x^2]}{\mathbb{E}[\tau]} + 2\nu \left(\frac{\mathbb{E}[\Delta x \tau]}{\mathbb{E}[\tau]} + \nu \frac{\mathbb{E}[\tau^2]}{\mathbb{E}[\tau]} \right)$$

  Applying the formula for $\hat{x}$ below (B.5), we have the result.

- **Reset state:** For the reset state $\hat{x}$, we apply Itō's Lemma to x_t^2 to obtain $d(x_t^2) = 2x_t dx_t + (dx_t)^2 = \left(2\nu x_t + \sigma^2\right) dt + 2\sigma x_t dW_t$. Using the OST $\mathbb{E}[\int_0^\tau x_s dW_s] = 0$. Moreover, given that $\mathbb{E}[\int_0^\tau x_s ds] = \mathbb{E}[x]\mathbb{E}[\tau] = 0$, we have that

$$\mathbb{E}[x_\tau^2] = \hat{x}^2 + \sigma^2 \mathbb{E}[\tau]. \tag{B.4}$$

Completing squares $\mathbb{E}[x_\tau^2] = \mathbb{E}[(\hat{x} - (\hat{x} - x_\tau))^2] = \mathbb{E}[\Delta x^2] - 2\hat{x}\mathbb{E}[\Delta x] + (\hat{x})^2$, we get

$$\begin{aligned}
\hat{x} &= \frac{1}{2\mathbb{E}[\Delta x]} \left[\mathbb{E}[\Delta x^2] - \sigma^2 \mathbb{E}[\tau] \right] \\
&= \frac{1}{2\mathbb{E}[\Delta x]} \left[\mathbb{E}[\Delta x^2] - \left(\mathbb{E}[\Delta x^2] + 2\frac{\mathbb{E}[\Delta x]\mathbb{E}[\Delta x \tau]}{\mathbb{E}[\tau]} + \frac{\mathbb{E}[\Delta x]^2 \mathbb{E}[\tau^2]}{\mathbb{E}[\tau]^2} \right) \right] \\
&= \frac{\mathbb{E}[\Delta x \tau]}{\mathbb{E}[\tau]} - \frac{\mathbb{E}[\Delta x]\mathbb{E}[\tau^2]}{2\mathbb{E}[\tau]^2}. \tag{B.5}
\end{aligned}$$

Rewriting in terms of covariance $\mathbb{C}ov[\tau, \Delta x] = \mathbb{E}[\tau \Delta x] + \mathbb{E}[\tau]\mathbb{E}[\Delta x]$ and coefficient of variation squared $\mathbb{CV}^2[\tau] = \frac{\mathbb{V}[\tau]}{\mathbb{E}[\tau]^2}$, we have the result.

- **Ergodic moments of x:** For observability of ergodic moments of x, apply Itō's Lemma to x^{m+1} and get $dx_t^{m+1} = (m+1)x_t^m \nu dt + (m+1)x_t^m \sigma dW_t + \frac{\sigma^2}{2}m(m+1)x_t^{m-1}dt$. Integrating from 0 to τ, using the OST to eliminate martingales, and rearranging:

$$\mathbb{E}\left[\int_0^\tau x_t^m dt\right] = \frac{1}{\nu(m+1)}\left(\mathbb{E}[x_\tau^{m+1}] - \hat{x}^{m+1}\right) - \frac{\sigma^2}{2\nu}m\mathbb{E}\left[\int_0^\tau x_t^{m-1}dt\right]. \tag{B.6}$$

Substituting the equivalences $\mathbb{E}[x^m] = \mathbb{E}\left[\int_0^\tau x_t^m dt\right]/\mathbb{E}[\tau]$ and $\mathbb{E}[\Delta x] = -\nu\mathbb{E}[\tau]$ yields:

$$\mathbb{E}[x^m] = \frac{\hat{x}^{m+1} - \mathbb{E}[(\hat{x} - \Delta x)^{m+1}]}{\mathbb{E}[\Delta x](m+1)} - \frac{\sigma^2 m}{2\nu}\mathbb{E}[x^{m-1}], \qquad \mathbb{E}[x] = 0. \tag{B.7}$$

- **Joint ergodic moments of x and age:** For observability of ergodic moments of $x^m a$, where a stands for the duration of the last action, we use Itō's Lemma and the OST on $x_t^{m+1}t$:

$$\mathbb{E}\left[\tau\left(\hat{x} - \Delta x\right)^{m+1}\right] = \mathbb{E}\left[\int_0^\tau x_t^{m+1}dt\right] + (m+1)\nu\mathbb{E}\left[\int_0^\tau x_t^m t dt\right] + \frac{\sigma^2 m(m+1)}{2}\mathbb{E}\left[\int_0^\tau x_t^{m-1}t dt\right] \tag{B.8}$$

and therefore

$$\mathbb{E}[x^m a] = \frac{\mathbb{E}\left[\tau\left(\hat{x} - \Delta x\right)^{m+1}\right]}{\nu(m+1)\mathbb{E}[\tau]} - \frac{\mathbb{E}[x^{m+1}]}{\nu(m+1)} - \frac{\sigma^2 m}{2\nu}\mathbb{E}[x^{m-1}a] \tag{B.9}$$

with initial condition $\mathbb{E}[a] = \frac{\mathbb{E}[\tau^2]}{2\mathbb{E}[\tau]}$.

$\square$

Proof of Proposition *2.* Here we extend the result in Álvarez, Le Bihan and Lippi (2016) for higher-order moments, arbitrary state, and policies. Start from the CIR's definition and fix an $m \in \mathbb{N}$:

$$\mathrm{CIR}_m = \mathbb{E}\left[\int_0^\infty \left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\right], \tag{B.10}$$

where the expectation is taken across agents ω. Using the strong Markov property of S and law of iterated expectations, we express the CIR as a function of the initial distribution F_0:

$$\mathrm{CIR}_m = \mathbb{E}\left[\mathbb{E}\left[\int_0^\infty \left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\,\middle|\,\mathcal{F}_0\right]\right] = \int_S \mathbb{E}\left[\int_0^\infty \left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\,\middle|\,S_0 = S\right]dF_0(S) =: \mathrm{CIR}_m(F_0), \quad \text{(B.11)}$$

Let $\{T_i\}_{i=1}^\infty$ be the sequence of stopping times. In equality (1) below, we write the CIR as the cumulative deviations between time $t = 0$ and the first stopping time T_1 plus the sum of deviations between all future stopping times. In (2), we use the Law of Iterated Expectations to condition on the information set $\mathcal{F}_{T_i}$. In (3), we use the strong Markov property of S_t, the assumption of homogeneous resets, and that $\hat{S}$ is constant for $i \geq 1$ to change the conditioning from $S_{T_i+h}|\mathcal{F}_{T_i}$ to $S_h|\hat{S}$ and write the problem recursively. We define $\tau = T_{i+1} - T_i$. In (4), we show that *every element* inside the infinite sum is equal to zero. For this purpose, recall the relationship between ergodic moments and expected duration derived in Auxiliary Theorem 2, $\mathbb{E}[x^m] = \mathbb{E}\left[\int_0^\tau x_t(\delta|\omega)^m|\hat{S}\right]/\mathbb{E}[\tau|\hat{S}]$, and thus we are left with the simple expression in the fourth line:

$$
\begin{aligned}
\mathrm{CIR}_m(F_0) \quad &= \quad \int_S \mathbb{E}\left[\int_0^\infty \left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\,\middle|\,S_0 = S\right]dF_0(S), &\text{(B.12)}\\[2mm]
&=^{(1)} \quad \int_S \mathbb{E}\left[\int_0^{T_1}\left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt + \sum_{i=1}^\infty \int_{T_i}^{T_{i+1}}\left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\,\middle|\,S_0 = S\right]dF_0(S)\\[2mm]
&=^{(2)} \quad \int_S \mathbb{E}\left[\int_0^{T_1}\left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt + \sum_{i=1}^\infty \mathbb{E}\left[\int_{T_i}^{T_{i+1}}\left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt\middle|\mathcal{F}_{T_i}\right]\,\middle|\,S_0 = S\right]dF_0(S)\\[2mm]
&=^{(3)} \quad \int_S \mathbb{E}\left[\int_0^{T_1}\left(x_t(\omega)^m - \mathbb{E}[x^m]\right)dt + \mathbb{E}\left[\sum_{i=1}^\infty \underbrace{\mathbb{E}\left[\int_0^\tau x_t(\omega)^m dt\middle|\hat{S}\right] - \mathbb{E}[x^m]\mathbb{E}[\tau|\hat{S}]}_{=0}\middle|\,S_0 = S\right]\right]dF_0(S)\\[2mm]
&=^{(4)} \quad \int_S \left[\mathbb{E}\left[\int_0^\tau \left(x_t(\omega)^m\right)dt\middle|S\right] - \mathbb{E}[x^m]\mathbb{E}[\tau|S]\,\middle|\,S_0 = S\right]dF_0(S)
\end{aligned}
$$

Note that we can write the CIR as follows, by defining the value function $\tilde{v}_m(S)$ conditional on a particular initial condition $S = [x, S^{-x}]$:

$$\text{CIR}_m(F_0) = \int_S \tilde{v}_m(S) dF_0(S), \qquad \tilde{v}_m(S) \equiv \mathbb{E}^S \left[\int_0^\tau (x_t(\omega)^m - \mathbb{E}[x^m]) \, dt \right] \tag{B.13}$$

Then rewrite the last expression as

$$
\begin{aligned}
\tilde{v}_m(S) &\equiv \mathbb{E}^S \left[\int_0^\tau x_t(\omega)^m dt \right] - \mathbb{E}^S \left[\int_0^\tau \mathbb{E}[x^m] dt \right] \\
&= \mathbb{E}^S \left[\int_0^\tau x_t(\omega)^m dt \right] - \mathbb{E}[x^m] \mathbb{E}^S \left[\int_0^\tau 1 dt \right] \\
&= \mathbb{E}^S \left[\int_0^\tau x_t(\omega)^m dt \right] - \mathbb{E}[x^m] \mathbb{E}^S \left[\int_0^\tau x_t(\omega)^0 dt \right] \\
&= v_m(S) - \mathbb{E}[x^m] v_0(S)
\end{aligned}
$$

where we define $v_m(S) \equiv \mathbb{E}^S \left[\int_0^\tau x_t^m dt \right]$. $\qquad\square$

Proof of Proposition 3. The idea behind this proof is that the CIR and the ergodic moments share the same finite basis $\Omega = \{e^{\xi_1 x}, e^{\xi_2 x}, e^{\xi_1 x} x, e^{\xi_2 x} x, \{x^i\}_{i=0}^m\}$, where ξ_1, ξ_2 are the characteristic roots of the HJB satisfied during periods of inaction. After this observation, the rest involves simple coefficient matching and algebra.

For clarity, we elaborate the proof in Lemmas 1 to 5.

(i) Lemma 1 rewrites the CIR, up to a first order approximation with respect to the aggregate shock δ, as the average of the marginal value of equation (36) given by $v'_m(x) - \mathbb{E}[x^m] v'_0(x)$. Then, we express the CIR as a combination of the following auxiliary functions: $\mathbb{E}\left[\int_0^\tau x_t^m dt \big| x \right]$, $\mathbb{E}\left[\int_0^\tau e^{\xi_1 x_t} dt \big| x \right]$, $\mathbb{E}\left[\int_0^\tau e^{\xi_2 x_t} dt \big| x \right]$ evaluated at $x = \hat{x}$.

(ii) Lemma 2 expresses the auxiliary functions from the previous Lemma in terms of the finite basis Ω.

(iii) Lemma 3 expresses $\mathbb{E}\left[\int_0^\tau t x_t^m dt \big| x \right]$ as a function of the same finite basis Ω. Note that this function is related to $\mathbb{E}[x^m a]$ through the occupancy measure.

(iv) Lemma 4 and 5 relate the basis coefficients between the CIR and the ergodic moments, obtaining two linear relationships that connect both objects:

$$\text{CIR}_m(\delta)/\delta = \frac{\mathbb{E}[x^{m+1}]}{\sigma^2} + \mathbb{L}(\{\mathbb{E}[x^i]\}_{i=0}^m) + \text{``discrepancy''}, \tag{B.14}$$

$$-\frac{\nu}{\sigma^2} \mathbb{C}ov[x^m, a] = \mathbb{L}(\{\mathbb{E}[x^i]\}_{i=0}^m) + \text{``discrepancy''}. \tag{B.15}$$

The first term in (B.14) relates to the ratio of the $m+1$ ergodic moment to idiosyncratic volatility $\mathbb{E}[x^{m+1}]/\sigma^2$; the second term is a linear combination of the first m ergodic moments $\mathbb{L}(\{\mathbb{E}[x^i]\}_{i=0}^m)$, where $\mathbb{L}(x)$ is a linear projection function; and the third term is a "discrepancy".

(v) Subtracting (B.15) from (B.14) we obtain the result:

$$\text{CIR}_m(\delta)/\delta = \frac{\mathbb{E}[x^{m+1}] - \nu \mathbb{C}ov[x^m, a]}{\sigma^2} \tag{B.16}$$

$\qquad\square$

Notation 1. *For the next Lemmas, we use the following notation:*

$$\tilde{\lambda}(\varphi) = \frac{\lambda - \varphi}{\sigma^2}$$

$$\xi_1(\varphi) = -\tilde{\nu} - \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}(\varphi)} \quad ; \quad \xi_2(\varphi) = -\tilde{\nu} + \sqrt{\tilde{\nu}^2 + 2\tilde{\lambda}(\varphi)} \quad ; \quad \xi_1(\varphi)\xi_2(\varphi) = 2\tilde{\lambda}(\varphi)$$

$$\mathcal{D}(\varphi) = e^{\xi_1(\varphi)\underline{x} + \xi_2(\varphi)\overline{x}} - e^{\xi_2(\varphi)\underline{x} + \xi_1(\varphi)\overline{x}}$$

$$\overline{\alpha}_1(\varphi) = \frac{e^{\xi_1(\varphi)\overline{x}}}{\mathcal{D}(\varphi)} \quad ; \quad \underline{\alpha}_1(\varphi) = \frac{e^{\xi_1(\varphi)\underline{x}}}{\mathcal{D}(\varphi)} \quad ; \quad \overline{\alpha}_2(\varphi) = \frac{e^{\xi_2(\varphi)\overline{x}}}{\mathcal{D}(\varphi)} \quad ; \quad \underline{\alpha}_2(\varphi) = \frac{e^{\xi_2(\varphi)\underline{x}}}{\mathcal{D}(\varphi)}$$

$$\kappa_m(x, \varphi) = \sum_{i=0}^m b_{i,m} x^i \quad \text{with} \quad b_{i,m}(\varphi) = \frac{m!}{i!} \left[\frac{\xi_1(\varphi) + 1/\phi}{\xi_1(\varphi) - \xi_2(\varphi)} \xi_1(\varphi)^{i-m} + \frac{\xi_2(\varphi) + 1/\phi}{\xi_2(\varphi) - \xi_1(\varphi)} \xi_2(\varphi)^{i-m} \right].$$

$$\underline{\kappa}_m(\varphi) = \kappa_m(\underline{x}, \varphi) \quad ; \quad \overline{\kappa}_m(\varphi) = \kappa_m(\overline{x}, \varphi) \quad ; \quad \hat{\kappa}_m(\varphi) = \kappa_m(\hat{x}, \varphi).$$

Moreover, we use the notation

$$\tilde{\nu} = \frac{\nu}{\sigma^2} \quad ; \quad \tilde{\lambda} = \tilde{\lambda}(0) \quad ; \quad \phi \equiv \sigma^2/2\nu$$

$$\xi_1 = \xi_1(0) \quad ; \quad \xi_2 = \xi_2(0) \quad ; \quad \mathcal{T} = \mathcal{T}(0)$$

$$\mathcal{D} = \mathcal{D}(0) \quad ; \quad \overline{\alpha}_1 = \overline{\alpha}_1(0) \quad ; \quad \underline{\alpha}_1 = \underline{\alpha}_1(0) \quad ; \quad \overline{\alpha}_2 = \overline{\alpha}_2(0) \quad ; \quad \underline{\alpha}_2 = \underline{\alpha}_2(0)$$

$$\mathcal{T} = (\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2)^{-1} \quad ; \quad b_{i,m} = b_{i,m}(0)$$

$$\kappa_m(x) \equiv \kappa_m(x,0) \quad ; \quad \overline{\kappa}_m = \kappa_m(\overline{x},0) \quad ; \quad \underline{\kappa}_m = \kappa_m(\underline{x},0) \quad ; \quad \hat{\kappa}_m = \kappa_m(\hat{x},0)$$

Lemma 1. *For this proof, we take $S = x$. Define $v_m(x) \equiv \mathbb{E}\left[\int_0^\tau x_t^m dt | x\right]$, then we can write the CIR as:*

$$CIR_m(\delta)/\delta = \frac{1}{\mathbb{E}[\tau]}\left(\mathbb{E}\left[\int_0^\tau v_m'(x_t)dt\right] - \mathbb{E}[x^m]\mathbb{E}\left[\int_0^\tau v_0'(x_t)dt\right]\right) + o(\delta) \tag{B.17}$$

with

$$\mathbb{E}\left[\int_0^\tau v_m'(x_t)dt\right] = \frac{1}{\lambda}\left(-\xi_1 V_1(\hat{x})\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right] - \xi_2 V_2(\hat{x})\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right] + \sum_{i=1}^m iv_{i-1}(\hat{x})b_{i,m}\right) \tag{B.18}$$

$$V_1(x) \equiv \mathbb{E}\left[\int_0^\tau e^{\xi_1 x_t}dt | x\right] \tag{B.19}$$

$$V_2(x) \equiv \mathbb{E}\left[\int_0^\tau e^{\xi_2 x_t}dt | x\right] \tag{B.20}$$

Proof. The random fixed adjustment cost model satisfies all of the assumptions in the preliminaries A. Thus

$$\text{CIR}_m(\delta) = \int_{\underline{x}}^{\overline{x}} \left(v_m(x) - \mathbb{E}[x^m]v_0(x)\right)f(x-\delta)dx$$

Doing a first-order Taylor approximation over δ (note that $f'(x)$ exist for all $x \in [\underline{x},\overline{x}]/\{\hat{x}\}$) and integrating by parts (using that there is no mass at the borders of inaction, i.e., $f(\underline{x}) = f(\overline{x}) = 0$), we have that

$$\text{CIR}_m(\delta) = -\delta \int_{\underline{x}}^{\overline{x}} \left(v_m(x) - \mathbb{E}[x^m]v_0(x)\right)f'(x)dx + o(\delta^2)$$

$$= -\delta\left[f(x)\left(v_m(x) - \mathbb{E}[x^m]v_0(x)\right)\Big|_{\underline{x}}^{\overline{x}} - \int_{\underline{x}}^{\overline{x}}\left(v_m'(x) - \mathbb{E}[x^m]v_0'(x)\right)dF(x)\right] + o(\delta^2)$$

$$= \delta \int_{\underline{x}}^{\overline{x}} \left(v_m'(x) - \mathbb{E}[x^m]v_0'(x)\right)dF(x) + o(\delta^2).$$

Now, we proceed to compute $v_m(x) \equiv \mathbb{E}\left[\int_0^\tau x_t^m dt \Big| x\right]$.

Computing $v_m(x)$. The function $v_m(x)$ satisfies the HJB

$$0 = x^m + \nu v_m'(x) + \frac{\sigma^2}{2}v_m''(x) - \lambda v_m(x)$$

with the border conditions $v_m(\underline{x}) = 0$ and $v_m(\overline{x}) = 0$. The homogeneous solution is given by

$$v_m^h(x) = B_{1,m}e^{\xi_1 x} + B_{2,m}e^{\xi_2 x}.$$

To find the non homogenous solution, we guess and verify a polynomial expression $v_m^c(x) = \sum_{i=0}^m \frac{b_{i,m}}{\lambda}x^i$. Plugged into the HJB

$$0 = \frac{x^m}{\sigma^2} + \tilde{\nu}\left(\sum_{i=1}^m \frac{b_{i,m}}{\lambda}ix^{i-1}\right) + \frac{1}{2}\left(\sum_{i=2}^m \frac{b_{i,m}}{\lambda}i(i-1)x^{i-2}\right) - \tilde{\lambda}\sum_{i=0}^m \frac{b_{i,m}}{\lambda}x^i.$$

For this equation to hold, the coefficients $\{b_{i,m}\}$ must satisfy the following difference equation:

$$b_{m,m} = 1, \qquad b_{m-1,m} = \frac{\nu}{\lambda}m, \qquad b_{i,m} = \frac{\nu}{\lambda}(i+1)\left[b_{i+1,m} + \phi b_{i+2,m}(i+2)\right]$$

The solution is given by

$$b_{i,m} = \frac{m!}{i!}\left[\frac{\xi_1 + 1/\phi}{\xi_1 - \xi_2}(\xi_1)^{i-m} + \frac{\xi_2 + 1/\phi}{\xi_2 - \xi_1}(\xi_2)^{i-m}\right].$$

The border conditions define a system of equations to be solved for $B_{1,m}$ and $B_{2,m}$

$$B_{1,m}e^{\xi_1\underline{x}} + B_{2,m}e^{\xi_2\underline{x}} = -\sum_{i=0}^{m}\frac{b_{i,m}}{\lambda}\underline{x}^i, \qquad B_{1,m}e^{\xi_1\overline{x}} + B_{2,m}e^{\xi_2\overline{x}} = -\sum_{i=0}^{m}\frac{b_{i,m}}{\lambda}\overline{x}^i.$$

Using Cramer's rule, we obtain (B.21).

$$
\begin{aligned}
B_{1,m} &= \frac{\left(-\sum_{i=0}^{m}b_{i,m}\underline{x}^i\right)e^{\xi_2\overline{x}} - \left(-\sum_{i=0}^{m}b_{i,m}\overline{x}^i\right)e^{\xi_2\underline{x}}}{\mathcal{D}} = -\left[\overline{\alpha}_2\underline{\kappa}^m - \underline{\alpha}_2\overline{\kappa}^m\right] \\
B_{2,m} &= \frac{\left(-\sum_{i=0}^{m}b_{i,m}\overline{x}^i\right)e^{\xi_1\underline{x}} - \left(-\sum_{i=0}^{m}b_{i,m}\underline{x}^i\right)e^{\xi_1\overline{x}}}{\mathcal{D}} = -\left[\underline{\alpha}_1\overline{\kappa}^m - \overline{\alpha}_1\underline{\kappa}^m\right].
\end{aligned}
$$

Thus we obtain the first result:

$$v_m(x) = \frac{e^{\xi_1 x}\left[\underline{\alpha}_2\overline{\kappa}_m - \overline{\alpha}_2\underline{\kappa}_m\right] + e^{\xi_2 x}\left[\overline{\alpha}_1\underline{\kappa}_m - \underline{\alpha}_1\overline{\kappa}_m\right] + \kappa_m(x)}{\lambda} \tag{B.21}$$

Writing the CIR as function of $V_1(\hat{x}), V_2(\hat{x})$. Using the relationship between ergodic moments and expected duration derived in Auxiliary Theorem 2,

$$\text{CIR}_m(\delta)/\delta = \int_{\underline{x}}^{\overline{x}}[v_m'(x) - \mathbb{E}[x^m]v_0'(x)]dF(x) + o(\delta) = \frac{\mathbb{E}\left[\int_0^\tau v_m'(x_t)dt\right] - \mathbb{E}[x^m]\mathbb{E}\left[\int_0^\tau v_0'(x_t)dt\right]}{\mathbb{E}[\tau]} + o(\delta).$$

Operating over $\mathbb{E}\left[\int_0^\tau v_m'(x_t)dt\right]$, we have that

$$\mathbb{E}\left[\int_0^\tau v_m'(x_t)dt\right] = \frac{1}{\lambda}\left(-\xi_1\mathbb{E}\left[\int_0^\tau e^{\xi_1 x_t}dt\right]\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right] - \xi_2\mathbb{E}\left[\int_0^\tau e^{\xi_2 x_t}dt\right]\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right] + \sum_{i=1}^{m}iv_{i-1}(\hat{x})b_{i,m}\right)$$

$\square$

Lemma 2. *The functions $V_1(x)$ and $V_2(x)$ can be expressed as*

$$V_1(x) = \frac{\mathcal{T}\left[e^{\xi_1 x}(\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}) + e^{\xi_2 x}\underline{\alpha}_1\overline{\alpha}_1(\overline{x} - \underline{x})\right] - e^{\xi_1 x}x}{\sigma^2(\tilde{\nu} + \xi_1)} \tag{B.22}$$

$$V_2(x) = \frac{\mathcal{T}\left[e^{\xi_2 x}(\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}) + e^{\xi_1 x}\underline{\alpha}_2\overline{\alpha}_2(\underline{x} - \overline{x})\right] - e^{\xi_2 x}x}{\sigma^2(\tilde{\nu} + \xi_2)}. \tag{B.23}$$

Proof. **Computing $V_1(x)$ and $V_2(x)$.** Next, we show (B.22) (computing (B.23) requires similar steps, mutatis mutandis). The value $V_1(x)$ satisfies the HJB

$$0 = e^{\xi_1 x} + \nu V_1'(x) + \frac{\sigma^2}{2}V_1''(x) - \lambda V_1(x)$$

with the border conditions $V_1(\underline{x}) = 0$ and $V_1(\overline{x}) = 0$. The homogeneous solution is given by

$$V_1^h(x) = A_1 e^{\xi_1 x} + A_2 e^{\xi_2 x}.$$

The non homogeneous solution satisfies $V_1^c(x) = e^{\xi_1 x}Cx$. Plugging it into the HJB we get

$$0 = e^{\xi_1 x} + \nu\left[e^{\xi_1 x}C + \xi_1 e^{\xi_1 x}Cx\right] + \frac{\sigma^2}{2}\left[2\xi_1 e^{\xi_1 x}C + \xi_1^2 e^{\xi_1 x}Cx\right] - \lambda e^{\xi_1 x}Cx,$$

Dividing everything by $\sigma^2 e^{\xi_1 x}$, we re-express it as

$$0 = \frac{1}{\sigma^2} + \tilde{\nu}\left[C + \xi_1 Cx\right] + \frac{1}{2}\left[2\xi_1 C + \xi_1^2 Cx\right] - \tilde{\lambda}Cx = x\underbrace{(\tilde{\nu}\xi_1 + \frac{\xi_1^2}{2} - \tilde{\lambda})}_{=0}C + \frac{1}{\sigma^2} + (\tilde{\nu} + \xi_1)C.$$

The previous equation must be satisfied for all x, thus $C = -\frac{1}{\sigma^2(\tilde{\nu}+\xi_1)}$. Using the border conditions, we find A_1 and A_2.

$$A_1 e^{\xi_1\underline{x}} + A_2 e^{\xi_2\underline{x}} = \frac{e^{\xi_1\underline{x}}\underline{x}}{\sigma^2(\tilde{\nu}+\xi_1)}, \qquad A_1 e^{\xi_1\overline{x}} + A_2 e^{\xi_2\overline{x}} = \frac{e^{\xi_1\overline{x}}\overline{x}}{\sigma^2(\tilde{\nu}+\xi_1)}$$

Using $e^{\xi_1\underline{x}} = \mathcal{T}\underline{\alpha}_1; e^{\xi_1\overline{x}} = \mathcal{T}\overline{\alpha}_1;\ e^{\xi_2\underline{x}} = \mathcal{T}\underline{\alpha}_2; e^{\xi_1\overline{x}} = \mathcal{T}\overline{\alpha}_2$, the solution is given by

$$V_1(x) = -\frac{\mathcal{T}\left[-e^{\xi_1 x}(\underline{\alpha}_1\underline{x}\overline{\alpha}_2 - \overline{\alpha}_1\overline{x}\underline{\alpha}_2) - e^{\xi_2 x}\underline{\alpha}_1\overline{\alpha}_1(\overline{x} - \underline{x})\right] + e^{\xi_1 x}x}{\sigma^2(\tilde{\nu}+\xi_1)},$$

Lemma 3. *Define the following value:*

$$U_m(x) \equiv \mathbb{E}\left[\int_0^\tau t x_t^m dt \Big| x\right],$$

It can be expressed as

$$\lambda U_m(x) = v_m(x) - e^{\xi_1 x}\left[\overline{\alpha}_2 \underline{\kappa}'_m(0) - \underline{\alpha}_2 \overline{\kappa}'_m(0)\right] - e^{\xi_2 x}\left[\underline{\alpha}_1 \overline{\kappa}_m(0) - \overline{\alpha}_1 \underline{\kappa}'_m(0)\right] + \kappa'_m(x,0)\ldots$$

$$\ldots + \frac{e^{\xi_1 x} x\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_1)} + \frac{e^{\xi_2 x} x\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_2)} - e^{\xi_1 x}\left[\overline{\alpha}'_2(0)\underline{\kappa}_m - \underline{\alpha}'_2(0)\overline{\kappa}_m\right] - e^{\xi_2 x}\left[\underline{\alpha}'_1(0)\overline{\kappa}_m - \overline{\alpha}'_1(0)\underline{\kappa}_m\right]$$

Proof. First we characterize $U_m(x)$ with the moment-generating function. Define the functions $H_m(x,\varphi) = \mathbb{E}^x[\int_0^\tau e^{\varphi t} x_t^m dt]$ and $\frac{H_m(x,\varphi)}{\partial \varphi}\Big|_{\varphi=0} = U_m(x)$. The function $H_m(x,\varphi)$ satisfies the HJB

$$-\varphi H_m(x,\varphi) = x^m + \nu \frac{\partial H_m(x,\varphi)}{\partial x} + \frac{\sigma^2}{2}\frac{\partial H_m(x,\varphi)}{\partial x^2} - \lambda H_m(x,\varphi)$$

with the border conditions $H_m(x,\varphi)(\underline{x},\varphi) = 0$ and $H_m(x,\varphi)(\overline{x},\varphi) = 0$. With similar steps as Lemma 2, we have

$$H_m(x,\varphi) = \frac{-e^{\xi_1(\varphi)x}\left[\overline{\alpha}_2(\varphi)\underline{\kappa}_m(\varphi) - \underline{\alpha}_2(\varphi)\overline{\kappa}_m(\varphi)\right] - e^{\xi_2 x}\left[\underline{\alpha}_1(\varphi)\overline{\kappa}_m(\varphi) - \overline{\alpha}_1(\varphi)\underline{\kappa}_m(\varphi)\right] + \hat{\kappa}_m(x,\varphi)}{\lambda - \varphi}.$$

Using the result that $\frac{d(e^{\xi_i(\varphi)\hat{x}})}{d\varphi} = -\frac{e^{\xi_i(\varphi)\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_i(\varphi))}$ for $i=1$ and $i=2$, we take the derivative of $h_m(\varphi)$ and evaluate it at $\varphi = 0$,

$$\lambda U_m(x) = v_m(x) - e^{\xi_1 x}\left[\overline{\alpha}_2 \underline{\kappa}'_m(0) - \underline{\alpha}_2 \overline{\kappa}'_m(0)\right] - e^{\xi_2 x}\left[\underline{\alpha}_1 \overline{\kappa}'_m(0) - \overline{\alpha}_1 \underline{\kappa}'_m(0)\right] + \kappa'_m(x,0)\ldots$$

$$\ldots + \frac{e^{\xi_1 x} x\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_1)} + \frac{e^{\xi_2 x} x\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_2)} - e^{\xi_1 x}\left[\overline{\alpha}'_2(0)\underline{\kappa}_m - \underline{\alpha}'_2(0)\overline{\kappa}_m\right] - e^{\xi_2 x}\left[\underline{\alpha}'_1(0)\overline{\kappa}_m - \overline{\alpha}'_1(0)\underline{\kappa}_m\right]$$

□

Lemma 4. *The following relationship holds*

$$CIR_m(\delta)/\delta \quad = \quad \frac{\mathbb{E}[x^{m+1}] - \frac{\nu}{\lambda}\sum_{i=0}^m b_{i,m}\mathbb{E}[x^i] + \mathbb{E}[x^m]\frac{\nu}{\lambda}}{\sigma^2} - \frac{\tilde{\nu}}{\lambda\sigma^2\mathbb{E}[\tau]}(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}) \tag{B.24}$$

$$\mathcal{K}_{2m} \quad = \quad \frac{\mathcal{L}_1}{\tilde{\nu}+\xi_1}\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right] + \frac{\mathcal{L}_2}{\tilde{\nu}+\xi_2}\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right] \tag{B.25}$$

$$\mathcal{L}_1 \quad = \quad \mathcal{T}\left[-e^{\xi_1 \hat{x}}(\underline{\alpha}_1 \underline{x}\overline{\alpha}_2 - \overline{\alpha}_1 \overline{x}\underline{\alpha}_2) - e^{\xi_2 \hat{x}}\underline{\alpha}_1 \overline{\alpha}_1(\overline{x} - \underline{x})\right] + e^{\xi_1 \hat{x}}\hat{x} \tag{B.26}$$

$$\mathcal{L}_2 \quad = \quad \mathcal{T}\left[-e^{\xi_1 \hat{x}}\underline{\alpha}_2 \overline{\alpha}_2(\underline{x} - \overline{x}) - e^{\xi_2 \hat{x}}(\overline{\alpha}_2 \overline{x}\underline{\alpha}_1 - \underline{\alpha}_2 \underline{x}\overline{\alpha}_1)\right] + e^{\xi_2 \hat{x}}\hat{x} \tag{B.27}$$

Proof. Departing from Lemma 1

$$\sigma^2 \lambda \mathbb{E}\left[\int_0^\tau v'_m(x_t) dt\right]$$

$$= -\xi_1 \sigma^2 V_1(\hat{x})\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right] - \xi_2 \sigma^2 V_2(\hat{x})\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right] + \sigma^2 \sum_{i=1}^m i v_{i-1}(\hat{x})b_{i,m}$$

$$= \frac{\xi_1 \mathcal{L}_1}{(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right] + \frac{\xi_2 \mathcal{L}_2}{(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right] + \sigma^2 \sum_{i=1}^m i v_{i-1}(\hat{x})b_{i,m}$$

$$= \underbrace{\mathcal{L}_1\left[\overline{\alpha}_2 \underline{\kappa}_m - \underline{\alpha}_2 \overline{\kappa}_m\right] + \mathcal{L}_2\left[\underline{\alpha}_1 \overline{\kappa}_m - \overline{\alpha}_1 \underline{\kappa}_m\right] + \sigma^2 \sum_{i=1}^m i v_{i-1}(\hat{x})b_{i,m}}_{\mathcal{K}_{1m}} - \tilde{\nu}\mathcal{K}_{2m},$$

where $\mathcal{K}_{2m}$ is defined in (B.25).

Departing from the definition of $\mathcal{K}_{1m}$ and using the property that $\mathcal{T}\left[\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2\right] = 1$ we have

$$\mathcal{K}_{1m} = \mathcal{L}_1\left[\overline{\alpha}_2\sum_{i=0}b_{i,m}\underline{x}^i - \underline{\alpha}_2\sum_{i=0}b_{i,m}\overline{x}^i\right] + \mathcal{L}_2\left[\underline{\alpha}_1\sum_{i=0}b_{i,m}\overline{x}^i - \overline{\alpha}_1\sum_{i=0}b_{i,m}\underline{x}^i\right] + \sigma^2\sum_{i=1}^{m}i\mathbb{E}\left[\int_0^\tau x^{i-1}\right]b_{i,m}$$

$$= \hat{x}[-v_m(\hat{x})\lambda + \hat{\kappa}_m] + \sigma^2\sum_{i=1}^{m}i\mathbb{E}\left[\int_0^\tau x^{i-1}\right]b_{i,m}\ldots$$

$$\ldots - e^{\xi_1\hat{x}}\mathcal{T}\left[\left[\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2\right]\left[\overline{\alpha}_2\sum_{i=0}^{m}b_{i,m}\underline{x}^{i+1} - \underline{\alpha}_2\sum_{i=0}^{m}b_{i,m}\overline{x}^{i+1}\right] + \sum_{i=0}^{m}b_{i,m}\underline{x}^i\overline{x}\underbrace{\left[\overline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2 - \overline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2\right]}_{=0} + \sum_{i=0}^{m}b_{i,m}\overline{x}^i\underline{x}\underbrace{\left[\underline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2 - \underline{\alpha}_1\overline{\alpha}_2\underline{\alpha}_2\right]}_{=0}\right]\ldots$$

$$\ldots - e^{\xi_2\hat{x}}\mathcal{T}\left[\left[\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2\right]\left[\underline{\alpha}_1\sum_{i=0}^{m}b_{i,m}\overline{x}^{i+1} - \overline{\alpha}_1\sum_{i=0}^{m}b_{i,m}\underline{x}^{i+1}\right] + \sum_{i=0}^{m}b_{i,m}\underline{x}^i\overline{x}\underbrace{\left[\overline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1 - \overline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1\right]}_{=0} + \sum_{i=0}^{m}b_{i,m}\overline{x}^i\underline{x}\underbrace{\left[\underline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1 - \underline{\alpha}_2\overline{\alpha}_1\underline{\alpha}_1\right]}_{=0}\right]$$

$$= \hat{x}[-v_m(\hat{x})\lambda + \hat{\kappa}_m] + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m}ib_{i,m}\left[-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\underline{\kappa}_{i-1} - \underline{\alpha}_2\overline{\kappa}_{i-1}\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\overline{\kappa}_{i-1} - \overline{\alpha}_1\underline{\kappa}_{i-1}\right] + \hat{\kappa}_{i-1}\right]\ldots$$

$$\ldots - e^{\xi_1\hat{x}}\left[\left[\overline{\alpha}_2\sum_{i=0}^{m}b_{i,m}\underline{x}^{i+1} - \underline{\alpha}_2\sum_{i=0}^{m}b_{i,m}\overline{x}^{i+1}\right]\right] - e^{\xi_2\hat{x}}\left[\left[\underline{\alpha}_1\sum_{i=0}^{m}b_{i,m}\overline{x}^{i+1} - \overline{\alpha}_1\sum_{i=0}^{m}b_{i,m}\underline{x}^{i+1}\right]\right].$$

$$= -\hat{x}\lambda v_m(\hat{x}) + \frac{1}{\tilde{\lambda}}\sum_{i=1}^{m}ib_{i,m}\sum_{j=0}^{i-1}b_{j,i-1}\Psi(j) + \sum_{i=0}^{m}b_{i,m}\Psi(i+1),$$

where we defined

$$\Psi(i) \equiv -e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\underline{x}^i - \underline{\alpha}_2\overline{x}^i\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\overline{x}^i - \overline{\alpha}_1\underline{x}^i\right] + \hat{x}^i.$$

In Auxiliary Result 1 we show that

$$\frac{1}{\tilde{\lambda}}\sum_{i=1}^{m}ib_{i,m}\sum_{j=0}^{i-1}b_{j,i-1}\Psi(j) + \sum_{i=0}^{m}b_{i,m}\Psi(i+1) = \sum_{i=0}^{m+1}b_{i,m+1}\Psi(i) - \frac{\nu}{\lambda}\sum_{i=0}^{m}b_{i,m}\sum_{j=0}^{i}b_{j,i}\Psi(j), \tag{B.28}$$

Finally, using (B.58), we have that

$$\mathcal{K}_{1,m} = \lambda\left(v_{m+1}(\hat{x}) - \hat{x}v_m(\hat{x}) - \frac{\nu}{\lambda}\sum_{i=0}^{m}b_{i,m}v_i(\hat{x})\right).$$

Note that

$$\mathcal{K}_{1,0} = \lambda\left(v_1(\hat{x}) - \hat{x}v_0(\hat{x}) - \frac{\nu}{\lambda}v_0(\hat{x})\right) = \lambda\mathbb{E}[\tau](-\hat{x}\frac{\nu}{\lambda}) \tag{B.29}$$

since $v_1(\hat{x}) = \mathbb{E}[x]\mathbb{E}[\tau] = 0$ and $v_0(\hat{x}) = \mathbb{E}[\tau]$. Using this result we have that

$$\begin{aligned}\mathrm{CIR}_m(\delta)/\delta &= \frac{\mathcal{K}_{1m} - \tilde{\nu}\mathcal{K}_{2m} - \mathbb{E}[x^m]\left(\mathcal{K}_{10} - \tilde{\nu}\mathcal{K}_{20}\right)}{\lambda\sigma^2\mathbb{E}[\tau]} + o(\delta)\\ &= \frac{v_{m+1}(\hat{x}) - \hat{x}v_m(\hat{x}) - \frac{\nu}{\lambda}\sum_{i=0}^{m}b_{i,m}v_i(\hat{x}) - \mathbb{E}[x^m]\mathbb{E}[\tau]\left(-\hat{x} - \frac{\nu}{\lambda}\right)}{\sigma^2\mathbb{E}[\tau]} - \frac{\tilde{\nu}}{\lambda\sigma^2\mathbb{E}[\tau]}\left(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}\right) + o(\delta)\\ &= \frac{\mathbb{E}[x^{m+1}] - \frac{\nu}{\lambda}\sum_{i=0}^{m}b_{i,m}\mathbb{E}[x^m] + \mathbb{E}[x^m]\frac{\nu}{\lambda}}{\sigma^2} - \frac{\tilde{\nu}}{\lambda\sigma^2\mathbb{E}[\tau]}\left(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}\right) + o(\delta)\end{aligned}$$

$\square$

Lemma 5. *The following relationships hold:*

$$-\frac{\nu}{\sigma^2}\mathbb{C}ov[x^m, a] = \frac{-\frac{\nu}{\lambda}\sum_{i=0}^{m}b_{i,m}\mathbb{E}[x^i] + \mathbb{E}[x^m]\frac{\nu}{\lambda}}{\sigma^2} - \frac{\tilde{\nu}}{\lambda\sigma^2\mathbb{E}[\tau]}\left(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}\right) \tag{B.30}$$

Proof. First we characterize $\mathbb{E}[ax^m]$ with the moment-generating function. From Lemma 3, we have

$$\lambda\mathbb{E}[ax^m]\mathbb{E}[\tau] = v_m(\hat{x}) \underbrace{-e^{\xi_1\hat{x}}\left[\overline{\alpha}_2\underline{\kappa}_m'(0) - \underline{\alpha}_2\overline{\kappa}_m'(0)\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1\overline{\kappa}_m'(0) - \overline{\alpha}_1\underline{\kappa}_m'(0)\right] + \hat{\kappa}_m'(0)}_{\mathcal{C}_{2m}}\ldots$$

$$\ldots + \underbrace{\frac{e^{\xi_1\hat{x}}\hat{x}\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_1)} + \frac{e^{\xi_2\hat{x}}\hat{x}\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_2)} - e^{\xi_1\hat{x}}\left[\overline{\alpha}_2'(0)\underline{\kappa}_m - \underline{\alpha}_2'(0)\overline{\kappa}_m\right] - e^{\xi_2\hat{x}}\left[\underline{\alpha}_1'(0)\overline{\kappa}_m - \overline{\alpha}_1'(0)\underline{\kappa}_m\right]}_{\mathcal{C}_{1m}}$$

(B.31)

Below, first we characterize $\mathcal{C}_{1m}$ and then $\mathcal{C}_{2m}$. Taking the derivative of $\underline{\alpha}_i(\varphi)$ and $\overline{\alpha}_i(\varphi)$, evaluating at $\varphi = 0$, and using the definition $\mathcal{T} = (\overline{\alpha}_2\underline{\alpha}_1 - \overline{\alpha}_1\underline{\alpha}_2)^{-1}$

$$\sigma^2\overline{\alpha}_2'(0) = \overline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\overline{\alpha}_1\underline{\alpha}_2(\overline{x}-\underline{x})}{\tilde{\nu}+\xi_2}\right] \quad ; \quad \sigma^2\underline{\alpha}_2'(0) = \underline{\alpha}_2\mathcal{T}\left[\frac{\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2(\overline{x}-\underline{x})}{\tilde{\nu}+\xi_2}\right] \quad \text{(B.32)}$$

$$\sigma^2\overline{\alpha}_1'(0) = \overline{\alpha}_1\mathcal{T}\left[\frac{-\underline{\alpha}_1\overline{\alpha}_2(\underline{x}-\overline{x})}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] \quad ; \quad \sigma^2\underline{\alpha}_1'(0) = \underline{\alpha}_1\mathcal{T}\left[\frac{-\overline{\alpha}_1\underline{\alpha}_2(\overline{x}-\underline{x})}{\tilde{\nu}+\xi_1} + \frac{\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}}{\tilde{\nu}+\xi_2}\right] \quad \text{(B.33)}$$

Using equations (B.32) and (B.33),

$$\mathcal{C}_{1m} = \frac{e^{\xi_1\hat{x}}\hat{x}\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_1)} + \frac{e^{\xi_2\hat{x}}\hat{x}\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right]}{\sigma^2(\tilde{\nu}+\xi_2)}\ldots$$

$$\ldots - \frac{e^{\xi_1\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}\right]\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right] - \frac{e^{\xi_1\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_2)}\overline{\alpha}_2\underline{\alpha}_2(\overline{x}-\underline{x})(\overline{\alpha}_1\underline{\kappa}_m - \underline{\alpha}_1\overline{\kappa}_m)\ldots$$

$$\ldots - \frac{e^{\xi_2\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_1)}\overline{\alpha}_1\underline{\alpha}_1(\underline{x}-\overline{x})\left[\underline{\alpha}_2\overline{\kappa}_m - \overline{\alpha}_2\underline{\kappa}_m\right] - \frac{e^{\xi_2\hat{x}}\mathcal{T}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}\right]\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right]$$

$$= \frac{-e^{\xi_1\hat{x}}\mathcal{T}\left[\underline{\alpha}_1\overline{\alpha}_2\underline{x} - \overline{\alpha}_1\underline{\alpha}_2\overline{x}\right] - e^{\xi_2\hat{x}}\mathcal{T}\overline{\alpha}_1\underline{\alpha}_1(\underline{x}-\overline{x}) + e^{\xi_1\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right]\ldots$$

$$\ldots + \frac{e^{\xi_1\hat{x}}\mathcal{T}\overline{\alpha}_2\underline{\alpha}_2(\overline{x}-\underline{x}) - e^{\xi_2\hat{x}}\mathcal{T}\left[\underline{\alpha}_1\overline{\alpha}_2\overline{x} - \overline{\alpha}_1\underline{\alpha}_2\underline{x}\right] + e^{\xi_2\hat{x}}\hat{x}}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right]$$

$$= \frac{\mathcal{L}_1}{\sigma^2(\tilde{\nu}+\xi_1)}\left[\overline{\alpha}_2\underline{\kappa}_m - \underline{\alpha}_2\overline{\kappa}_m\right] + \frac{\mathcal{L}_2}{\sigma^2(\tilde{\nu}+\xi_2)}\left[\underline{\alpha}_1\overline{\kappa}_m - \overline{\alpha}_1\underline{\kappa}_m\right]$$

$$= \frac{\mathcal{K}_{2m}}{\sigma^2},$$

therefore we have characterized $\mathcal{C}_{1m}$. Using the definition of $\Psi(i)$ in equation (B.28), we can write $\mathcal{C}_{2m}$ as

$$\mathcal{C}_{2m} = \sum_{i=0}^{m-1} b_{i,m}'(0)\Psi(i) \tag{B.34}$$

In Auxiliary Result 2, we show that

$$b_{i,m}'(0) = \frac{1}{\lambda}\sum_{j=i}^{m-1} b_{i,j}b_{j,m} \ \forall i \leq m-1, \tag{B.35}$$

Using equations (B.34) and (B.71), we have that

$$\mathcal{C}_{2m} = \sum_{i=0}^{m-1} b_{i,m}'(0)\Psi(i) = \sum_{i=0}^{m-1}\frac{1}{\lambda}\sum_{j=i}^{m-1} b_{i,j}b_{j,m}\Psi(i) = \frac{1}{\lambda}\sum_{j=0}^{m-1}\sum_{i=0}^{j} b_{j,m}b_{i,j}\Psi(i) = \frac{1}{\lambda}\sum_{j=0}^{m-1} b_{j,m}v_j(\hat{x})\lambda,$$

therefore we have characterized $\mathcal{C}_{2m}$.

We can now proof (B.30). Departing from equation (B.31)

$$\mathbb{E}[ax^m] = \frac{v_m'(\hat{x}) + \mathcal{C}_{1m} + \mathcal{C}_{2m}}{\mathbb{E}[\tau]\lambda}$$

$$= \frac{\mathbb{E}[x^m] + \mathcal{K}_{2m}/(\sigma^2\mathbb{E}[\tau]) + \sum_{i=0}^{m-1} b_{i,m}\mathbb{E}[x^i]}{\lambda}$$

$$= \frac{\mathcal{K}_{2m}/(\sigma^2\mathbb{E}[\tau]) + \sum_{i=0}^{m} b_{i,m}\mathbb{E}[x^i]}{\lambda} \tag{B.36}$$

Applying the formula for the covariance and using equation (B.36)

$$
\begin{aligned}
-\frac{\nu}{\sigma^2}\mathbb{C}ov[x^m, a] &= -\frac{\nu}{\sigma^2}\left(\mathbb{E}[ax^m] - \mathbb{E}[a]\mathbb{E}[x^m]\right) \\
&= -\frac{\nu}{\sigma^2}\left(\frac{\mathcal{K}_{2m}/(\sigma^2\mathbb{E}\left[\tau\right]) + \sum_{i=0}^m b_{i,m}\mathbb{E}[x^i]}{\lambda} - \frac{\mathcal{K}_{20}/(\sigma^2\mathbb{E}\left[\tau\right]) + 1}{\lambda}\mathbb{E}[x^m]\right) \\
&= \frac{-\frac{\nu}{\lambda}\sum_{i=0}^m b_{i,m}\mathbb{E}[x^m] + \mathbb{E}[x^m]\frac{\nu}{\lambda}}{\sigma^2} - \frac{\nu}{\lambda\sigma^4\mathbb{E}\left[\tau\right]}\left(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}\right) \\
&= \frac{-\frac{\nu}{\lambda}\sum_{i=0}^m b_{i,m}\mathbb{E}[x^m] + \mathbb{E}[x^m]\frac{\nu}{\lambda}}{\sigma^2} - \frac{\tilde{\nu}}{\lambda\sigma^2\mathbb{E}\left[\tau\right]}\left(\mathcal{K}_{2m} - \mathbb{E}[x^m]\mathcal{K}_{20}\right)
\end{aligned}
$$

$\square$

Subtracting (B.30) from (B.24) we finish the proof.

Proof of Proposition 4. The proof is a special case of Proposition 3. $\qquad\square$

Proposition 5. *Assume that the uncontrolled state follows $d\tilde{x}_t = \nu dt + \sigma dW_t$, where W_t is a Wiener process. Let $\phi \equiv \sigma^2/2\nu$. Define the following auxiliary function:*

$$
g_k(x) \;\equiv\; \mathbb{E}\left[x_\tau^k\Big|\hat{x} + x\right] - \mathbb{E}^{\hat{x}}\left[(x_\tau + x)^k\Big|\hat{x}\right]. \tag{B.37}
$$

Then, to a first order, the CIR_m is given by

$$
CIR_m(\delta)/\delta = \sum_{k=0}^m \frac{m!}{k!}(-\phi)^{m-k}\left(\Gamma_k + \Psi_k\right) - \mathbb{E}[x^m]\Psi_0 + o(\delta) \tag{B.38}
$$

where the margins of adjustment are given by

$$
\Gamma_k \;=\; k\mathbb{E}[x^{k-1}a] + \mathbb{1}_{\{k\geq 2\}}k(k-1)\phi\mathbb{E}[x^{k-2}a] \tag{B.39}
$$

$$
\Psi_k \;=\; \sum_{j=0}^\infty \psi_{k,j}\mathbb{E}[x^j] \tag{B.40}
$$

$$
\psi_{k,j} \;=\; \frac{1}{\nu}\sum_{i\geq j}^\infty \frac{\hat{x}^{i-j}}{i!j!}\left[\frac{d^{i+1}g_{k+1}(0)}{dx^{i+1}}/(k+1) - \frac{d^i g_k(0)}{dx^i}\right]. \tag{B.41}
$$

Proof. The proof is divided into three Lemmas for clarity. $\qquad\square$

Lemma 6. *[Aggregation]* *To a first order, the CIR_m is given by*

$$
CIR_m(\delta)/\delta = \sum_{k=0}^m (-\phi)^k \frac{m!}{(m-k)!}\left(\Gamma_{m-k} + \Psi_{m-k}\right) - \mathbb{E}[x^m]\Psi_0 + o(\delta) \tag{B.42}
$$

where the intensive and extensive margins are given by

$$
\Gamma_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi_m^\Gamma(x_t)dt\right]}{\mathbb{E}[\tau]} \quad;\quad \varphi_m^\Gamma(x_t) \equiv \frac{1}{\nu}\left(\mathbb{E}^x\left[x_\tau^m\right] - x_t^m\right) \tag{B.43}
$$

$$
\Psi_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi_m^\Psi(x_t)dt\right]}{\mathbb{E}[\tau]} \quad;\quad \varphi_m^\Psi(S_t) \equiv \frac{1}{\nu}\left[\frac{\partial\mathbb{E}^x\left[x_\tau^{m+1}/(m+1)\right]}{\partial x} - \mathbb{E}^x\left[x_\tau^m\right]\right] \tag{B.44}
$$

Proof. Throughout this proof we denote conditional expectations with superscripts, as follows $\mathbb{E}^y[\cdot] = \mathbb{E}[\cdot|y]$.

Express the CIR as derivatives with respect to initial condition. Start from the representation of the CIR in (B.17) from Lemma 1

$$
CIR_m(\delta)/\delta = \int \left(v_m'(x) - \mathbb{E}[x^m]v_0'(x)\right)dF(x) + o(\delta) \tag{B.45}
$$

where $v_m(x) \equiv \mathbb{E}^x\left[\int_0^\tau x_t^m dt\right]$ and $v_0(x) \equiv \mathbb{E}^x\left[\tau\right]$. From v_m's definition, take its derivative with respect to initial conditions and substitute it back into the CIR

$$
CIR_m(\delta)/\delta \;=\; \int \frac{\partial}{\partial x}\mathbb{E}^x\left[\int_0^\tau x_t^m dt\right]dF(x) - \mathbb{E}[x^m]\int \frac{\partial\mathbb{E}^S[\tau]}{\partial x}dF(x)
$$

49

Lastly, by adding and subtracting the following object $\int \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] dF(x)$, we re-express the CIR as the sum of three terms $\mathcal{B}_m$, $\mathcal{C}_m$, and Ψ_0, defined in the brackets.

$$\text{CIR}_m(\delta)/\delta = \underbrace{\int \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] dF(x)}_{\mathcal{B}_m} - \underbrace{\mathbb{E}[x^m] \int \frac{\partial \mathbb{E}^x[\tau]}{\partial x} dF(x)}_{\Psi_0} + \underbrace{\int \left(\frac{\partial}{\partial x} \mathbb{E}^x \left[\int_0^\tau x_t^m dt \right] - \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] \right) dF(x)}_{\mathcal{C}_m}.$$

(B.46)

Now we further characterize each of these terms.

Characterize $\mathcal{B}_m$. Since $x_t = x + \nu t + \sigma W_t$, for all $t \leq \tau$ we have that

$$\mathcal{B}_m \equiv \int \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] dF(x) = \int \mathbb{E}^x \left[\int_0^\tau m x_t^{m-1} dt \right] dF(x).$$

Applying Itō's Lemma to x_t^m we have $dx_t^m = \nu m x_t^{m-1} dt + \sigma m x_t^{m-1} dW_t + \frac{\sigma^2}{2} m(m-1) x_t^{m-2} dt$, and integrating both sides from 0 to τ and taking expectations with initial condition x, we get

$$\mathbb{E}^x \left[x_\tau^m \right] - x^m = m\sigma \underbrace{\mathbb{E}^x \left[\int_0^\tau x_t^{m-1} dW_t \right]}_{=0 \text{ by OST}} + \nu \mathbb{E}^x \left[\int_0^\tau m x_t^{m-1} dt \right] + \frac{\sigma^2 m}{2} \mathbb{E}^x \left[\int_0^\tau (m-1) x_t^{m-2} dt \right].$$

Given that $\int_0^t x_s^m dW_t$ is a martingale with zero initial condition and it is well defined by assumption, we apply the Optional Sampling Theorem (OST) to conclude that $\mathbb{E}^x \left[\int_0^\tau x_t^{m-1} dW_t \right] = 0$. Solve for $\mathbb{E}^x \left[\int_0^\tau m x_t^{m-1} dt \right]$:

$$\mathbb{E}^x \left[\int_0^\tau m x_t^{m-1} dt \right] = \frac{\mathbb{E}^x \left[x_\tau^m \right] - x^m}{\nu} - m\phi \mathbb{E}^x \left[\int_0^\tau (m-1) x_t^{m-2} dt \right].$$

Integrating both sides across all initial conditions, defining $\varphi_m^\Gamma \equiv \frac{1}{\nu} \left(\mathbb{E}^x \left[x_\tau^m \right] - x^m \right)$ and $\Gamma_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi_m^\Gamma(x_t) dt \right]}{\mathbb{E}[\tau]}$, and recognizing $\mathcal{B}_m$ and $\mathcal{B}_{m-1}$, we get:

$$\mathcal{B}_m = \Gamma_m - m\phi \mathcal{B}_{m-1}, \qquad \Gamma_0 = 0,$$ (B.47)

where we used the Auxiliary Theorem 2, exchanging the ergodic distribution for the local occupancy measure.

Characterize $\mathcal{C}_m$. With similar steps as in the previous point, we characterize $\mathcal{C}_m$ as follows:

$$\mathcal{C}_m \equiv \int \underbrace{\frac{\partial}{\partial x} \mathbb{E}^x \left[\int_0^\tau x_t^m dt \right]}_{A} - \underbrace{\mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right]}_{B} dF(x).$$

First, we get an expression for the term A. Applying Itō's Lemma to x_t^{m+1} we have $dx_t^{m+1} = (m+1)\nu x_t^m dt + \sigma(m+1) x_t^m dW_t + \frac{\sigma^2}{2} m(m+1) x_t^{m-1} dt$. Integrating both sides from 0 to τ, taking expectations with initial condition x, using the OST, and rearranging, we get $\mathbb{E}^x \left[\int_0^\tau x_t^m dt \right] = \frac{1}{\nu(m+1)} \left(\mathbb{E}^x \left[x_\tau^{m+1} \right] - x^{m+1} \right) - m\phi \mathbb{E}^x \left[\int_0^\tau x_t^{m-1} dt \right]$, and its derivative with respect to initial condition x:

$$A \equiv \frac{\partial}{\partial x} \mathbb{E}^x \left[\int_0^\tau x_t^m dt \right] = \frac{1}{\nu} \left(\frac{\partial \mathbb{E}^x \left[x_\tau^{m+1} \right] / (m+1)}{\partial x} - x_t^m \right) - m\phi \frac{\partial \mathbb{E}^x \left[\int_0^\tau x_t^{m-1} dt \right]}{\partial x}$$

Now, for the term B, recall from the characterization of Γ_m that

$$B \equiv \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] = \frac{1}{\nu} \left(\mathbb{E}^x \left[x_\tau^m \right] - x_t^m \right) - m\phi \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^{m-1}}{\partial x} dt \right].$$

Subtract the equations for A and B and simplify to obtain

$$\frac{\partial \mathbb{E}^x \left[\int_0^\tau x_t^m dt \right]}{\partial x} - \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^m}{\partial x} dt \right] = \frac{1}{\nu} \left(\frac{\partial \mathbb{E}^x \left[x_\tau^{m+1} \right] / (m+1)}{\partial x} - \mathbb{E}^x \left[x_\tau^m \right] \right) - m\phi \left\{ \frac{\partial \mathbb{E}^x \left[\int_0^\tau x_t^{m-1} dt \right]}{\partial x} - \mathbb{E}^x \left[\int_0^\tau \frac{\partial x_t^{m-1}}{\partial x} dt \right] \right\}$$

Integrate with the ergodic distribution, define $\varphi_m^\Psi(x_t) \equiv \frac{1}{\nu} \left(\frac{\partial \mathbb{E}^x \left[x_\tau^{m+1} \right] / (m+1)}{\partial x} - \mathbb{E}^x \left[x_\tau^m \right] \right)$ and $\Psi_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi_m^\Psi(x_t) dt \right]}{\mathbb{E}[\tau]}$, and recognize $\mathcal{C}_m$ and $\mathcal{C}_{m-1}$ to get:

$$\mathcal{C}_m = \Psi_m - m\phi \mathcal{C}_{m-1}, \qquad \mathcal{C}_{-1} = 0.$$ (B.48)

Characterize Ψ_0. We corroborate that the expression $\int \frac{\partial \mathbb{E}^x[\tau]}{\partial x} dF(x)$ is equal to Ψ_0. By the OST, we have $\mathbb{E}^x[x_\tau] - x = \nu \mathbb{E}^x[\tau]$. Thus $\frac{\partial \mathbb{E}^x[\tau]}{\partial x} = \frac{1}{\nu} \left[\frac{\partial \mathbb{E}^x[x_\tau]}{\partial x} - 1 \right]$. Substituting and using Auxiliary Theorem 2, we recover the expression for Ψ_0 in the

50

definition of Ψ_m:

$$\Psi_0 \equiv \int \frac{\partial \mathbb{E}^x[\tau]}{\partial x} dF(x) = \int \frac{1}{\nu} \left[\frac{\partial \mathbb{E}^x[x_\tau]}{\partial x} - 1 \right] dF(x)$$

All results together. Define $\mathcal{Z}_m \equiv \mathcal{B}_m + \mathcal{C}_m$, which implies $\mathcal{Z}_m = \Gamma_m + \Psi_m - m\phi\mathcal{Z}_{m-1}$. Combine the results in (B.46), (B.47) and (B.48) to obtain $\text{CIR}_m(\delta)/\delta = (\mathcal{Z}_m - \mathbb{E}[x^m]\Psi_0)$. Note that we can write $\mathcal{Z}_m$:

$$
\begin{aligned}
\mathcal{Z}_m &= \Gamma_m + \Psi_m + (-\phi)^1 m \mathcal{Z}_{m-1} \\
&= \Gamma_m + \Psi_m + (-\phi)^1 m \left(\Gamma_{m-1} + \Psi_{m-1} \right) + (-\phi)^2 m(m-1)\mathcal{Z}_{m-2} \\
&= \Gamma_m + \Psi_m + (-\phi)^1 m \left(\Gamma_{m-1} + \Psi_{m-1} \right) + (-\phi)^2 m(m-1) \left(\Gamma_{m-2} + \Psi_{m-2} \right) + (-\phi)^3 m(m-1)(m-2)\mathcal{Z}_{m-3} \\
&= \ldots \\
&= \sum_{k=1}^{m} (-\phi)^{m-k} \frac{m!}{k!} \left(\Gamma_k + \Psi_k \right)
\end{aligned}
$$

$\square$

Lemma 7. *[Representation for intensive margin]* The intensive margin Γ_m defined as

$$\Gamma_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi^\Gamma(x_t) dt \right]}{\mathbb{E}[\tau]}, \qquad with \qquad \varphi_m^\Gamma(x_t) = \frac{1}{\nu} \left(\mathbb{E}^x[x_\tau^m] - x_t^m \right).$$

can be represented as a function of steady-state moments as

$$\Gamma_m = m\mathbb{E}[x^{m-1}a] + \mathbb{1}_{\{m \geq 2\}} \phi m(m-1) \mathbb{E}[x^{m-2}, a].$$

Proof. Start (1) from the definition of Γ_m and $\varphi_m^\Gamma(S)$, then (2) exchange the time integral with the expectation conditional on adjustment $\mathbb{E}[\cdot]$, which introduces an indicator $\mathbb{1}_{\{t \leq \tau\}}$. Use the Law of Iterated Expectations in (3) to condition on the set $\{t \leq \tau\}$.

$$
\begin{aligned}
\nu\mathbb{E}[\tau]\Gamma_m &=^{(1)} \mathbb{E}\left[\int_0^\tau \mathbb{E}^{x_t}[x_\tau^m] - x_t^m dt \right] =^{(2)} \int_0^\infty \mathbb{E}\left[\left(\mathbb{E}^{x_t}[x_\tau^m] - x_t^m \right) \mathbb{1}_{\{t \leq \tau\}} \right] dt \\
&=^{(3)} \int_0^\infty \mathbb{E}\left[\mathbb{E}\left\{ \left(\mathbb{E}^{x_t}[x_\tau^m] - x_t^m \right) \mathbb{1}_{\{t \leq \tau\}} \Big| t \leq \tau \right\} \right] dt =^{(4)} \int_0^\infty \mathbb{E}\left[\mathbb{E}\left\{ \mathbb{E}^{x_t}[x_\tau^m] - x_t^m \Big| t \leq \tau \right\} \right] dt \\
&=^{(5)} \int_0^\infty \mathbb{E}\left[\mathbb{E}\left[x_\tau^m - x_t^m \Big| t \leq \tau \right] \right] dt =^{(6)} \int_0^\infty \mathbb{E}\left[\mathbb{E}\left\{ \left(x_\tau^m - x_t^m \right) \mathbb{1}_{\{t \leq \tau\}} \Big| t \leq \tau \right\} \right] dt \\
&=^{(7)} \mathbb{E}\left[\int_0^\tau \left(x_\tau^m - x_t^m \right) dt \right] =^{(8)} \mathbb{E}\left[x_\tau^m \int_0^\tau dt \right] - \mathbb{E}\left[\int_0^\tau x_t^m dt \right]
\end{aligned}
\tag{B.49}
$$

We now characterize $\mathbb{E}\left[x_\tau^{m+1} \int_0^\tau dt \right]$. Applying Itō's Lemma followed by the OST to $Y_t^m \equiv x_t^m \int_0^t ds$

$$dY_t^m = x_t^m dt + \nu m x_t^{m-1} \int_0^t ds dt + \sigma m x_t^{m-1} \int_0^t ds dW_t + \mathbb{1}_{\{m \geq 2\}} \frac{m(m-1)\sigma^2}{2} x_t^{m-2} \int_0^t ds dt \tag{B.50}$$

$$\mathbb{E}[Y_\tau^m] = \mathbb{E}\left[x_\tau^m \int_0^\tau dt \right] = \underbrace{\mathbb{E}\left[\int_0^\tau x_t^m dt \right]}_{\mathbb{E}[x^m]\mathbb{E}[\tau]} + \nu m \underbrace{\mathbb{E}\left[\int_0^\tau x_t^{m-1} t dt \right]}_{\mathbb{E}[x^{m-1}a]\mathbb{E}[\tau]} + \mathbb{1}_{\{m \geq 2\}} \frac{m(m-1)\sigma^2}{2} \underbrace{\mathbb{E}\left[\int_0^\tau x_t^{m-2} t dt \right]}_{\mathbb{E}[x^{m-2}a]\mathbb{E}[\tau]} \tag{B.51}$$

Using equations (B.49) and (B.51), we have that

$$\Gamma_m = m\mathbb{E}[x^{m-1}a] + \mathbb{1}_{\{m \geq 2\}} \phi m(m-1) \mathbb{E}[x^{m-2}a].$$

$\square$

Lemma 8. *[Representation for extensive margin]* Assume the moments of the adjustment size can be written as

$$g_m(x) = \mathbb{E}^{\hat{x}+x}[(\hat{x} - \Delta x)^m] - \mathbb{E}[(\hat{x} - \Delta x + x)^m]. \tag{B.52}$$

Then the extensive margin given by $\Psi_m \equiv \frac{\mathbb{E}\left[\int_0^\tau \varphi_m^\Psi(x_t) dt \right]}{\mathbb{E}[\tau]}$ with $\varphi_m^\Psi(x) \equiv \frac{1}{\nu} \left(\frac{\partial \mathbb{E}^x[x_\tau^{m+1}/(m+1)]}{\partial x} - \mathbb{E}^x[x_\tau^m] \right)$ can be represented as a function of steady-state moments as follows:

$$\Psi_m = \sum_{j=0}^\infty \psi_{m,j} \mathbb{E}[x^j] \quad with \quad \theta_{m,j} = \frac{1}{\nu} \sum_{k \geq j}^\infty \frac{\hat{x}^{k-j}}{k!j!} \left[\frac{d^{k+1}g_{m+1}(0)}{dx^{k+1}} / m + 1 - \frac{d^k g_m(0)}{dx^k} \right]. \tag{B.53}$$

Proof. Using a change of variable in assumption (B.52), we have that

$$\mathbb{E}^y\left[x_\tau^m\right] = g_m(y-\hat{x}) + \mathbb{E}\left[(y-\hat{x}+x_\tau)^m\right]. \tag{B.54}$$

Using the previous equation, we have that

$$\Psi_m = \frac{\mathbb{E}\left[\int_0^\tau \left(\frac{\partial \mathbb{E}^x[x_\tau^{m+1}/(m+1)]}{\partial x} - \mathbb{E}^x[x_\tau^m]\right)dt\right]}{\nu\mathbb{E}\left[\tau\right]} \tag{B.55}$$

$$= \frac{1}{\nu}\frac{\mathbb{E}\left[\int_0^\tau \left[\left[\frac{dg_{m+1}(y-\hat{x})}{dy}/(m+1) + \mathbb{E}\left[(y-\hat{x}+x_\tau)^{m+1}\right]\right] - \left[g_m(y-\hat{x}) + \mathbb{E}\left[(y-\hat{x}+x_\tau)^m\right]\right]\right]dt\right]}{\mathbb{E}[\tau]}$$

$$= \frac{1}{\nu}\frac{\mathbb{E}\left[\int_0^\tau \left[\frac{dg_{m+1}(y-\hat{x})}{dy}/(m+1) - g_m(y-\hat{x})\right]dt\right]}{\mathbb{E}[\tau]}$$

$$= \frac{1}{\nu}\frac{\mathbb{E}\left[\int_0^\tau \left[\sum_{j=0}^\infty \frac{d^j}{j!dx^j}\left[\frac{dg_{m+1}(x)}{dx}/(m+1) - g_m(x)\right]_{x=0}(y-\hat{x})^j\right]dt\right]}{\mathbb{E}[\tau]}$$

$$= \frac{1}{\nu}\sum_{j=0}^\infty\sum_{z=0}^j \frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{j!}\left(\begin{array}{c}j\\z\end{array}\right)\frac{\mathbb{E}\left[\int_0^\tau x_t^{j-z}dt\right]}{\mathbb{E}[\tau]}$$

$$= \frac{1}{\nu}\sum_{j=0}^\infty\sum_{z=0}^j \frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{j!}\left(\begin{array}{c}j\\z\end{array}\right)\mathbb{E}[x^{j-z}]$$

$$= \sum_{j=0}^\infty\sum_{z=0}^j \underbrace{\frac{1}{\nu}\frac{\left[\frac{d^{j+1}g_{m+1}(0)}{dx^{j+1}}/m+1 - \frac{d^j g_m(0)}{dx^j}\right]\hat{x}^z}{z!(j-z)!}}_{=\mathcal{H}_{m,j,z}}\mathbb{E}[x^{j-z}]$$

$$= \sum_{h=0}^\infty \psi_{m,h}\mathbb{E}[x^h], \ with \ \psi_{m,h} = \sum_{k\geq h}^\infty \mathcal{H}_{m,k,k-h}. \tag{B.56}$$

If $\tau|x_t \sim \tau$ (as in fully time-dependent models), then we have that

$$\begin{aligned}
g_m(x) &= \mathbb{E}^{\hat{x}+x}\left[(\hat{x}-\Delta x)^m\right] - \mathbb{E}\left[(\hat{x}-\Delta x+x)^m\right]\\
&= \mathbb{E}\left[(\hat{x}+x+\nu\tau^{x+\hat{x}}+\sigma W_{\tau^{x+\hat{x}}})^m\right] - \mathbb{E}\left[(\hat{x}+x+\nu\tau+\sigma W_\tau)^m\right]\\
&= \mathbb{E}\left[(\hat{x}+x+\nu\tau+\sigma W_\tau)^m\right] - \mathbb{E}\left[(\hat{x}+x+\nu\tau+\sigma W_\tau)^m\right]\\
&= 0
\end{aligned} \tag{B.57}$$

$\square$

Auxiliary Result 1. *[Equivalence in coefficients]*

$$\frac{1}{\tilde{\lambda}}\sum_{i=1}^m ib_{i,m}\sum_{j=0}^{i-1}b_{j,i-1}\Psi(j) + \sum_{i=0}^m b_{i,m}\Psi(i+1) = \sum_{i=0}^{m+1}b_{i,m+1}\Psi(i) - \frac{\nu}{\lambda}\sum_{i=0}^m b_{i,m}\sum_{j=0}^i b_{j,i}\Psi(j), \tag{B.58}$$

Proof. Now, we show (B.58). It is easy to verify the following in difference equations for the $\{\{b_{i,m}\}_{i=0}^m\}$:

$$b_{m,m} = 1 \quad ; \quad b_{m-1,m} = m\frac{\nu}{\lambda} \tag{B.59}$$

$$b_{i,m} = \frac{\nu}{\lambda}(i+1)b_{i+1,m} + \frac{1}{\tilde{\lambda}2}(i+2)(i+1)b_{i+2,m} \quad with \ i \leq m-2 \tag{B.60}$$

$$b_{i,m} = b_{i-1,m-1}\frac{m}{i} \quad with \ i \geq 1 \ and \ m \geq 1 \tag{B.61}$$

$$b_{i,m} = \frac{\nu}{\lambda}mb_{i,m-1} + \frac{(m-1)m}{\tilde{\lambda}2}b_{i,m-2} \quad with \ i \leq m-2 \ \ m \geq 2. \tag{B.62}$$

Operate over the left-hand side of (B.58) and we have

$$LHS = \sum_{j=0}^{m-1} \sum_{i=j+1}^{m} \frac{ib_{i,m}b_{j,i-1}}{\tilde{\lambda}} \Psi(j) + \sum_{j=1}^{m} b_{j-1,m} \Psi(j) + b_{m,m} \Psi(m+1)$$

$$= \sum_{i=1}^{m} \frac{ib_{i,m}b_{0,i-1}}{\tilde{\lambda}} \Psi(0) + \sum_{j=1}^{m-1} \left[\sum_{i=j+1}^{m} \frac{ib_{j,i-1}b_{i,m}}{\tilde{\lambda}} + b_{j-1,m} \right] \Psi(j) + b_{m-1,m} \Psi(m) + b_{m,m} \Psi(m+1), \qquad (B.63)$$

and operating over the right-hand side of (B.58)

$$RHS = \sum_{j=0}^{m} b_{j,m+1} \Psi(j) - \sum_{j=0}^{m} \sum_{i=j}^{m} \frac{\nu}{\lambda} b_{i,m}b_{j,i} \Psi(j) + b_{m+1,m+1} \Psi(m+1)$$

$$= \sum_{j=0}^{m} \left[b_{j,m+1} - \frac{\nu}{\lambda} \sum_{i=j}^{m} b_{j,i}b_{i,m} \right] \Psi(j) + b_{m+1,m+1} \Psi(m+1). \qquad (B.64)$$

Thus, to show that (B.63) is equal to (B.64), we need to get that for all m

$$b_{m+1,m+1} = b_{m,m} \qquad (B.65)$$

$$b_{m,m+1} - \frac{\nu}{\lambda} \sum_{i=m}^{m} b_{m,i}b_{i,m} = b_{m-1,m} \qquad (B.66)$$

$$b_{j,m+1} - \frac{\nu}{\lambda} \sum_{i=j}^{m} b_{j,i}b_{i,m} = \sum_{i=j+1}^{m} \frac{ib_{j,i-1}b_{i,m}}{\tilde{\lambda}} + b_{j-1,m} \quad with \quad m-1 \geq j \geq 1 \qquad (B.67)$$

$$b_{0,m+1} - \frac{\nu}{\lambda} \sum_{i=0}^{m} b_{0,i}b_{i,m} = \sum_{i=1}^{m} \frac{ib_{0,i-1}b_{i,m}}{\tilde{\lambda}} \qquad (B.68)$$

Now we show equations (B.65) to (B.68) by induction hypothesis. We skip the verification for $m = 1$ and $m = 2$. Assume they hold for $m-1$ and $m-2$, and now we show they hold for m:

- Equation (B.65): From equation (B.59) we have that $b_{m,m} = 1 = b_{m+1,m+1}$.

- Equation (B.66): From equation (B.59) we have that

$$b_{m-1,m} = m\frac{\nu}{\lambda} = \frac{\nu}{\lambda}(m+1) - \frac{\nu}{\lambda} = \frac{\nu}{\lambda}(m+1) - \frac{\nu}{\lambda} b_{m,m}b_{m,m} = b_{m,m+1} - \frac{\nu}{\lambda} \sum_{i=m}^{m} b_{m,i}b_{i,m} \qquad (B.69)$$

- Equation (B.67): Departing from the right-hand side of (B.67), we have that

$$= \sum_{i=j+1}^{m} \frac{ib_{j,i-1}b_{i,m}}{\tilde{\lambda}} + b_{j-1,m}$$

$$=^{(1)} \frac{(m-1)b_{j,m-2}b_{m-1,m}}{\tilde{\lambda}} + \frac{mb_{j,m-1}b_{m,m}}{\tilde{\lambda}} + \frac{m\nu}{\lambda} \left[\sum_{i=j+1}^{m-2} \frac{ib_{j,i-1}b_{i,m-1}}{\tilde{\lambda}} + b_{j-1,m-1} \right] + \frac{(m-1)m}{2\tilde{\lambda}} \left[\sum_{i=j+1}^{m-2} \frac{ib_{j,i-1}b_{i,m-2}}{\tilde{\lambda}} + b_{j-1,m-2} \right]$$

$$=^{(2)} \frac{(m-1)b_{j,m-2}b_{m-1,m}}{\tilde{\lambda}} + \frac{mb_{j,m-1}b_{m,m}}{\tilde{\lambda}} - \frac{m\nu}{\lambda} \frac{(m-1)b_{j,m-2}b_{m-1,m-1}}{\tilde{\lambda}} \cdots$$

$$\cdots + \frac{m\nu}{\lambda} \left[b_{j,m} - \frac{\nu}{\lambda} \sum_{i=j}^{m-1} b_{j,i}b_{i,m-1} \right] + \frac{(m-1)m}{2\tilde{\lambda}} \left[b_{j,m-1} - \frac{\nu}{\lambda} \sum_{i=j}^{m-2} b_{j,i}b_{i,m-2} \right]$$

$$=^{(3)} \frac{mb_{j,m-1}b_{m,m}}{\tilde{\lambda}} - \frac{m\nu}{\lambda} \frac{\nu}{\lambda} b_{j,m-1}b_{m-1,m-1} + \frac{m\nu}{\lambda} \left[b_{j-1,m} - \frac{\nu}{\lambda} \sum_{i=j}^{m-2} b_{j,i}b_{i,m-1} \right] + \frac{(m-1)m}{2\tilde{\lambda}} \left[b_{j,m-1} - \frac{\nu}{\lambda} \sum_{i=j}^{m-2} b_{j,i}b_{i,m-2} \right]$$

$$=^{(4)} -\frac{m\nu}{\lambda} \frac{\nu}{\lambda} b_{j,m-1}b_{m-1,m-1} + \frac{m\nu}{\lambda} b_{j,m} + \frac{(m+1)m}{2\tilde{\lambda}} b_{j,m-1} - \frac{\nu}{\lambda} \sum_{i=j}^{m-2} b_{j,i}b_{i,m}$$

$$=^{(5)} -\frac{\nu}{\lambda} b_{j,m} - \frac{\nu}{\lambda} b_{j,m-1}b_{m-1,m} + b_{j,m+1} - \frac{\nu}{\lambda} \sum_{i=j}^{m-2} b_{j,i}b_{i,m}$$

$$=^{(6)} b_{j,m+1} - \frac{\nu}{\lambda} \sum_{i=j}^{m} b_{j,i}b_{i,m} \qquad (B.70)$$

53

Now we explain the steps in equation (B.70). In step (1) we apply (B.62), and in step (2) we apply the induction hypothesis. In step (3) we operate, and in steps (4) and (5) we apply (B.62), (B.59) and operate. Step (6) completes the sum.

- Equation (B.68): With similar steps as before

$$b_{0,m+1} = \frac{\nu}{\lambda}(m+1)b_{0,m} + \frac{(m+1)m}{\tilde{\lambda}2}b_{0,m-1}$$

$$= \frac{\nu}{\lambda}(m+1)\left[\sum_{i=1}^{m-1}\frac{ib_{0,i-1}b_{i,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-1}b_{0,i}b_{i,m-1}\right] + \frac{(m+1)m}{\tilde{\lambda}2}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-2}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-2}\right]$$

$$= \frac{\nu}{\lambda}(m+1)\left[\frac{(m-1)b_{0,m-2}b_{m-1,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}b_{0,m-1}b_{m-1,m-1}\right]\cdots$$

$$\cdots + \frac{\nu}{\lambda}(m+1)\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{j,i}b_{i,m-1}\right] + \frac{(m+1)m}{\tilde{\lambda}2}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-2}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{j,i}b_{i,m-2}\right]$$

$$= \frac{\nu}{\lambda}(m+1)\left[\frac{(m-1)b_{0,m-2}b_{m-1,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}b_{0,m-1}b_{m-1,m-1}\right] + \frac{\nu}{\lambda}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-1}\right]\cdots$$

$$\cdots + \frac{m}{\tilde{\lambda}}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-2}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-2}\right]$$

$$\cdots + \frac{\nu}{\lambda}m\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-1}\right] + \frac{(m-1)m}{\tilde{\lambda}2}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-2}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{j,i}b_{i,m-2}\right]$$

$$= \left[\frac{(m-1)b_{0,m-2}b_{m-1,m}}{\tilde{\lambda}} + \frac{\nu}{\lambda}b_{0,m-1}b_{m-1,m}\right] + \frac{\nu}{\lambda}\left[\sum_{i=1}^{m-1}\frac{ib_{0,i-1}b_{i,m-1}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-1}\right]\cdots$$

$$\cdots + \frac{m}{\tilde{\lambda}}\left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m-2}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m-2}\right] + \left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-2}b_{0,i}b_{i,m}\right]$$

$$= \frac{\nu}{\lambda}b_{0,m}b_{m,m} + \frac{m}{\tilde{\lambda}}b_{0,m-1}b_{m,m} + \left[\sum_{i=1}^{m-2}\frac{ib_{0,i-1}b_{i,m}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m-1}b_{0,i}b_{i,m}\right]$$

$$= \sum_{i=1}^{m}\frac{ib_{0,i-1}b_{i,m}}{\tilde{\lambda}} + \frac{\nu}{\lambda}\sum_{i=0}^{m}b_{0,i}b_{i,m}$$

$\square$

Auxiliary Result 2.

$$b'_{i,m}(0) = \frac{1}{\lambda}\sum_{j=i}^{m-1}b_{i,j}b_{j,m} \ \forall i \le m-1, \tag{B.71}$$

Proof. We use the induction hypothesis of order two for any m over i. For this step, we need the following recursive characterization of $b_{i,m}$ and $b'_{i,m}(0)$, which is easy to show:

$$b_{m,m} = 1 \ ; \ b_{m-1,m} = \frac{\nu}{\lambda}m \ ; \ b_{i,m} = \frac{\nu}{\lambda}(i+1)b_{i+1,m} + \frac{(i+2)(i+1)}{2\tilde{\lambda}}b_{i+2,m} \ \forall i \le m-2. \tag{B.72}$$

$$b'_{m-1,m}(0) = \frac{1}{\lambda}\frac{\nu}{\lambda}m \ ; \ b'_{i,m}(0) = \frac{\nu}{\lambda}(i+1)b'_{i+1,m}(0) + \frac{(i+2)(i+1)}{2\tilde{\lambda}}b'_{i+2,m}(0) + \frac{b_{i,m}}{\lambda} \ \forall i \le m-2. \tag{B.73}$$

First, we verify that (B.71) holds for $i = m-1$ and $i = m-2$.

$$b'_{m-1,m}(0) = \frac{\nu m}{\lambda^2} = \frac{b_{m-1,m-1}b_{m-1,m}}{\lambda}$$

$$b'_{m-2,m}(0) = \frac{2\left(\frac{\nu}{\lambda}\right)^2(m-1)m + \frac{\nu}{2\lambda\lambda}}{\lambda} = \frac{b_{m-2,m-2}(0)b_{m-2,m}(0) + b_{m-2,m-1}(0)b_{m-1,m}(0)}{\lambda}.$$

54

Fix any $i \leq m - 2$ and assume that equation (B.71) holds for $i + 1$ and $i + 2$. Below, we equation (B.71) for i

$$
\begin{aligned}
b'_{i,m}(0) =^{(1)} \ & \frac{\nu}{\lambda}(i+1)b'_{i+1,m}(0) + \frac{(i+2)(i+1)}{2\tilde{\lambda}}b'_{i+2,m}(0) + \frac{b_{i,m}}{\lambda} \\
=^{(2)} \ & \frac{\nu}{\lambda}(i+1)\frac{1}{\lambda}\sum_{j=i+1}^{m-1}b_{i+1,j}b_{j,m} + \frac{(i+2)(i+1)}{2\tilde{\lambda}}\frac{1}{\lambda}\sum_{j=i+2}^{m-1}b_{i+2,j}b_{j,m} + \frac{b_{i,m}}{\lambda} \\
=^{(3)} \ & \frac{b_{i,m} + \frac{\nu}{\lambda}(i+1)b_{i+1,i+1}b_{i+1,m} + \sum_{j=i+2}^{m-1}\left[\frac{\nu}{\lambda}(i+1)b_{i^*+1,j} + \frac{(i+2)(i^*+1)b_{i+2,j}}{2\tilde{\lambda}}\right]b_{j,m}}{\lambda} \\
=^{(4)} \ & \frac{b_{i,m} + \frac{\nu}{\lambda}(i+1)b_{i+1,i+1}b_{i+1,m} + \sum_{j=i+2}^{m-1}b_{i,j}b_{j,m}}{\lambda} \\
=^{(5)} \ & \frac{b_{i,i}b_{i,m} + b_{i,i+1}b_{i+1,m} + \sum_{j=i+2}^{m-1}b_{i,j}b_{j,m}}{\lambda} \\
= \ & \frac{\sum_{j=i}^{m-1}b_{i,j}b_{j,m}}{\lambda}.
\end{aligned}
\tag{B.74}
$$

Now we explain each step in equation (B.74). In step (1) we use (B.73), and in step (2) we use the induction hypothesis. Step (3) operates over the first sum, and step (4) uses (B.73). Finally, step (5) uses that $b_{i,i} = 1$ and $\frac{\nu}{\lambda}(i+1)b_{i+1,i+1} = b_{i,i+1}$. Thus, we prove (B.71). $\qquad\square$