University of Glasgow

Adam Smith Business School

WORKING PAPER SERIES

Signal-jamming in the Frequency Domain

Bart Taub

Paper No. 2023-02
January 2023

SIGNAL-JAMMING IN THE FREQUENCY DOMAIN

B. TAUB*

Abstract. I examine strategic behaviour for a duopoly in a noisy environment. Firms attempt to learn the value of the rival's privately observed demand shocks via a noisy signal of price, and at the same time firms attempt to obfuscate that signal by producing excess output on the publicly observable signals, that is, they signal jam.

In a dynamic setting firms also distort the intertemporal structure of output keyed to the publicly observable demand shock process in order to disguise their private shocks. The net outcome is to radically increase the persistence of output over its full-information value.

JEL Classification Numbers: C61, C62, C63, C73, D43, D82.

Keywords: Dynamic games, signal jamming, strategic information, frequency-domain methods.

April 12, 2022

*University of Glasgow.

This project is descended from a joint project with Dan Bernhardt. I acknowledge his many contributions to the development of this paper. Much of this research was carried out in part during my stay at ICEF, Higher School of Economics, Moscow; I acknowledge financial support from the Academic Excellence Project "5-100" of the Russian Government. Additional research was carried out at the European University Institute, Florence, with the support of a Fernand Braudel Fellowship.

1. Introduction

Firms are continually buffeted over time by shocks to demand, some of which they learn about through direct common observation, some through private observation, and some of which they attempt to extract from information in price signals indirectly from their rivals' private knowledge. In this paper I examine what happens in a *dynamic* setting where firms keep learning over time about past shocks, and internalize how their actions affect a rival's information and future actions. I thus take on a longstanding challenge set by Mirman, Samuelson and Urbano [42] who observe that "the most appropriate model [is] an infinite horizon model in which the parameters of demand curves are subject to continual shocks. Firms are then repeatedly forced to draw inferences about unknown demand curves and to consider the effects of their actions on their rival's beliefs."

The structure of the model is as follows. There are two firms. Demand evolves according to a stationary and persistent autoregressive stochastic process with three independent components: a publicly-observed component, and two additional components each of which is observed privately by each of the two firms in the duopoly.[1] The firms also observe price, but only via a noisy signal, with the noise shock process common to both firms. Each firm combines the information extracted from the history of price signals with that in the history of its privately- and publicly-observed demand shocks to determine how much to produce, separately for each demand shock. A key element of the model is that the underlying persistence of the public and privately observed demand shocks (as indexed by the autoregressive parameters) can be different.

Because the fundamental demand shock processes are stationary, the equilibrium output strategies are stationary linear functions of the history of private signals and prices. The solution determines these functions and the resulting output processes by characterizing their coefficients, i.e., the *weights* put on shocks and prices.These linear functions are then the focus of the analysis. Each firm's profit maximisation problem can then be expressed as a variational problem in the so-called frequency domain, in which the firms choose these linear functions. The game between the firms can then be re-posed in the space of these functions, with the equilibrium a fixed point in the space.

The resulting equilibrium stochastic processes of output can have persistence properties that are radically different from the exogenous demand shock processes, due entirely to the strategic incentives to interfere with the information extraction of the rival firms, that is, to signal jam.

Related literature. The analysis of oligopolistic competition in supply schedules with demand uncertainty dates back to Klemperer and Meyer [36]. They argue that competition in supply schedules better describes strategic competition between firms than competition in prices or quantities, because it more realistically allows firms to adjust to market conditions. In particular, in a supply schedule equilibrium, firms adjust to market conditions in an optimal manner given their rival's behavior—given knowledge of the market-clearing price,

[1] Although it is possible to include cost shocks in the model, I eschew this dimension for notational simplicity.

they have no incentive to adjust outputs. In contrast, with stochastic Bertrand or Cournot competition, firms would want to alter their actions after learning something about demand.

Bernhardt and Taub [13] analyze the static counterpart to the dynamic model in this paper. That model is related to Vives [59]. In his static setting, firms receive private noisy signals about costs, and costs are correlated across firms, so one firm's signal is relevant for a rival. Firms compete in supply schedules and there is no demand uncertainty. As a result, the market-clearing price is privately fully revealing: in equilibrium, a firm's own cost signal and price yield the same forecast of its costs as when a firm also sees each rival's cost signals. [2]

An early signal-jamming literature explores belief manipulation in two-date models, where firms are symmetrically uninformed about demand or costs and learn from prices (Riordan [44] , Aghion et al. [2], [3], Mirman, Samuelson and Urbano [42], Caminal and Vives [18], Harrington [27], Alepuz and Urbano [4]). Firms condition date-2 output on date-1 price, inducing firms to over-produce at date 1 to lower price to try to persuade rivals that the market is less profitable. With no private information, firms perfectly learn in equilibrium at date 2.[3]

Keller and Rady [34] analyze symmetric learning in a continuous-time duopoly setting in which demand evolves according to a two-state Markov process and a firm *perfectly observes* its rival's actions.[4] By contrast, in this setting, the learning process is entangled with the strategic efforts of firms to manipulate the beliefs of rivals. Bergin and Bernhardt [12] analytically characterize the stationary entry and exit dynamics of a competitive industry when both common value demand and individual firm costs evolve according to Markov processes.

A large literature analyzes collusion with imperfect monitoring and common unobserved public shocks (e.g., Green and Porter [24], Abreu et al. [1], Sannikov [49], Hackbarth and Taub [25]). In these dynamic models, actions by rivals are unobserved, but are perfectly inferred in equilibrium because firms have the incentive to follow equilibrium "recommended" actions, and this means that punishments can be exacted for the failure to implement the recommended actions; this threat structure then supports the equilibrium. Similarly, with privately-observed costs, Athey and Bagwell [6] analyze collusion in a procurement auction game in which a firm's costs evolve according to a two-state Markov process, and firms make cheap-talk announcements about costs before bidding. Histories matter for incentives, but, with cheap talk, are not used to glean information about fundamentals. In contrast, in the model of this paper, inferences about a rival's privately-observed fundamentals are obscured by noise; because actions cannot be directly inferred it is not possible to threaten direct punishments. The equilibrium therefore rests on the strategic interaction between learning about primitives from prices and belief manipulation.

[2]Bergemann et al. [11] also analyze a static model with learning in which agents learn from private signals and prices.

[3]There is also a literature in which firms have private information about demand or costs, and take actions (e.g., limit price) to signal it. See Harrington [27], [28], Caminal [17], Bagwell and Ramey [7], or Mailath [39].

[4]Foundational papers on learning and experimentation by a monopolist include McLennan [41], Aghion et al. [2](1991), Harrington [29], Rustichini and Wolinsky [48] and Keller and Rady [34].

The model shares similarities of information and equilibrium structure with the financial speculation models descended from the model of Kyle [37] and [38]. In these models, informed traders interact with uninformed market makers who observe a noisy signal of the informed trades. The market makers extract information from that signal using Kalman filtering to determine price; the informed trader understands this and shades his trades accordingly to husband his information. In [37], the informed trader understands the net impact of his trades on price—he submits a demand schedule—and the market maker similarly understands that the price determination is simultaneous with this.

Similarly, the model here assumes that the firms possess private information about a fraction of the demand shocks and choose their output simultaneous with that observation. But in addition, just as in the Kyle model, they understand the impact of their output on price and therefore net out that impact from the noisy signal of price; thus there is no delay between the observation of price and the determination of output in response to the *net* information in price; all actions and observations are simultaneous.

This approach is also used by Bonatti, Cisterna and Toikka [16]. In their continuous-time, finite-horizon model, firms receive private-value cost shocks at the outset, Brownian motion demand shocks shift the equilibrium price, and firms learn about a rival's costs via price histories. As in this model, firms strategically manipulate price signals by overproducing.

A key feature of Kyle's model and its descendants is that informed traders *disguise* their trades so that when their trading orders are combined with the those of the noise traders, the resulting total flow of trade is indistiguishable from the noise trade with a higher variance, thus preventing market makers from inverting the total order flow to infer the informed traders' private information; this is known as "inconspicuousness." In the model here the firms behave similarly: the output process on public shocks has the same dynamic character—autoregressive structure—as the output on the private shock processes, thus preventing a rival from inferring the private demand shock process by inverting the price signal.

Frequency-domain methods. In the usual time-domain approach, each period, given the history of signals, a firm's period output function maximizes expected profits given correct beliefs about a rival's past and future optimization. Due to the model's linear-quadratic, Gaussian, time-separable, stationary structure, optimal policy rules are linear weightings of information histories. Along an equilibrium path these weights *do not change*: they are independent of the realized shock history, and hence remain optimal in the future: along an equilibrium path, optimal strategies are stationary. Frequency-domain methods provide an algebraic approach to determine these strategies.

Frequency domain applications in the literature. Hansen and Sargent [26] noted that the first-order conditions stemming from models in which expectation of future endogenous variables could be z-transformed, i.e., mapped to the frequency domain, and solved; the idea is akin to Fourier transforming a function. Whiteman [60], building on the work of Davenport and Root [20], saw that the optimization problem itself could be z-transformed and the optimization expressed as a variational problem and solved in the frequency domain.

Appendix A sets out these techniques and also validates the equivalence of the method with conventional time-domain optimization.

This paper also relates to research on the "forecasting the forecasts of others" endogenous information problem. Several papers attack the problem using frequency domain methods. Kasa [33] uses frequency-domain methods to show that the forecasting problem alone, which arises in rational expectations models with atomistic agents, simplifies in the frequency domain, as the infinite regress that would appear in the time domain collapses to a single function in the frequency domain. Kasa, Walker and Whiteman [32] also model information heterogeneity in an infinite-horizon, infinite history setting, taking advantage of the ability of frequency domain method to handle this complication. Rondina and Walker [45] model an endogenous information equilibrium problem, characterizing the equilibrium signals as a non-invertible reduced-form matrix of functions in the frequency domain. Makarov and Rytchkov [40] study a model with small, risk-averse investors who behave competitively, showing there is no finite representation of the equilibrium: because the endogenous variables have infinitely many poles, the equilibrium time domain processes cannot have Markovian dynamics—equivalently, finitely many state variables. Huo and Takayama [31] also find that if information is endogenous, an infinite regress problem develops and no finite representation is possible. These results echo similar findings in Seiler and Taub [51] who also demonstrate an infinite poles result.

Nimark [43] looks at endogenous information aggregation in a linear rational expectations model. He iterates on the Euler equation from a representative agent's optimization problem in a setting with endogenous variables such as prices, accounting for the dependence of those variables on the solution to the Euler equation. Using Hilbert space methods, he derives a contraction property to obtain the equilibrium. Seiler and Taub [51] and Bernhardt, et al. [14] carry out an analogous iteration in the frequency domain, leading to a contraction property.

Plan of the paper. The plan of the paper is as follows. In the next section I set out a recapitulation of the static model of Bernhardt and Taub [13], but with an exploration of some additional facets of the model that serve as intuitive benchmarks for the dynamic model. In section 3 I set up and provide the solutions for equilibrium firm behaviour. I then characterise the dynamic behaviour of output and prices in the dynamic setting in section 4. In section 5 I illustrate the characterisations with numerical simulations for three canonical examples.

Following the conclusion there are seven appendices. Three of these appendices are pedagogical in nature: Appendix A describes the frequency-domain methods used here in greater detail, Appendix F describes the state-space numerical methods that must be used to simulate the model, and Appendix G illustrates the algebra of spectral densities that underlies the simulation results. The other appendices, Appendices B-E, contain derivations and proofs for the substantive elements of the paper; existence is demonstrated using a fixed point argument for the appropriate space of functions in Appendix D.

2. A benchmark static model

To develop intuition I set out an abbreviated version of the static duopoly model in Bernhardt and Taub [13]. There is a homogeneous-good duopoly, where firms 1 and 2 face stochastic demand, which is the sum of three mean-zero independent shocks: $\overline{a}$ with variance $\sigma^2_{\overline{a}}$, which is publicly observable, a_1, and a_2, each with variance σ^2_a, and which are privately observable by firms 1 and 2 respectively. Firms cannot observe price directly; instead they see a noisy signal of price, with a common noise source, $p + e$, where $e \sim N(0, \sigma^2_e)$ is an independently-distributed unobservable demand shock. I normalize production costs to zero.[5] Define the vector of shocks by

$$X \equiv (\overline{a}, a_1, a_2, e). \tag{1}$$

Price is determined by these shocks and output:

$$p = \overline{a} + a_1 + a_2 - (q_1 + q_2) \tag{2}$$

However firms cannot observe price directly; instead they see a noisy signal of price, with a common noise source:

$$p + e = \overline{a} + a_1 + a_2 - (q_1 + q_2) + e \tag{3}$$

where $e \sim N(0, \sigma^2_e)$ is an independently-distributed unobservable demand shock.

The firms simultaneously observe the signal of price, which determines their output, whilst at the same time influencing the price with their output. However the firms are aware of their influence on price, and they net out the influence of their own output on price and then react to the net information in the price signal to determine output. Thus, output decisions and their effect on the price are partially decoupled, just as they would be if price were observable with a lag. The model can therefore be viewed as allowing firms to obtain a noisy signal of the rival's privately observed demand shock via the noisy price signal.[6] To maintain the spirit of the model I assume that firms cannot see realized profit without noise either, and because of this they cannot invert profit to infer the price.[7]

Firms compete in supply schedules. Because the model is linear, this means that rather than choosing output in response to price directly, they choose a linear coefficient on the price signal, which I term the *intensity*, which then determines output. Firm i's supply schedule, $Q^i : (p + e; \overline{a}, a_i) \mapsto \mathbb{R}$, is a differentiable function that maps each possible price signal, and demand shocks $\overline{a}$ and a_i into an output level. A price function $\pi : (q_1, q_2, X) \mapsto \mathbb{R}$ is *market clearing* if it is consistent with the supply schedules. Thus, for each X,

$$q_1 = Q^1(p + e; \overline{a}, a_1), \quad q_2 = Q^2(p + e; \overline{a}, a_2), \quad \text{and} \quad p = \pi(q_1, q_2, X). \tag{4}$$

This implicitly defines a fixed point problem in the space of functions containing π, Q^1, and Q^2.

---

[5]This is not a benign assumption, as production costs are private-value in character, in contrast to the common-value nature of demand shocks. As was demonstrated in Bernhardt and Taub [13], the equiibrium response to these shocks is then fundamentally different from the response to common-value shocks.

[6]Thus, as discussed in the introduction, the information structure is similar to that in the papers of Kyle [37] and Bonatti, Cisterna and Toikka [16].

[7]For example, ongoing inflation can add noise to the value of money in the calculation of profit.

I next verify the existence of an equilibrium in which supply functions are linear, taking the form

$$Q^i(p;\overline{a},a_i) = \alpha_i a_i + \beta_i \bar{a} + \delta_i(p+e). \tag{5}$$

Definition 1. *A linear equilibrium is a pair of supply functions $Q^{i*}(\cdot;\cdot,\cdot)$ with weights $(\alpha_i,\beta_i,\delta_i)$ satisfying (5) for $i \in \{1,2\}$, and a market-clearing price function $\pi^{P*}(\cdot,\cdot;\cdot)$ such that for each $(\overline{a},a_1,a_2,e)$: $Q^{i*}(p^{P*};\overline{a},a_i)$ maximizes firm i's expected profit given $\{p^{P*},\bar{a},a_i\}$; and prices and output are market clearing.*

To construct a linear equilibrium, substitute the conjectured linear form for a rival firm's supply function into the market-clearing price,

$$p = \overline{a} + a_1 + a_2 - q_i - (\alpha_{-i}a_{-i} + \beta_{-i}\bar{a} + \delta_{-i}(p+e)),$$

yielding the noisy price signal,

$$p + e = \frac{a_i + (1-\alpha_{-i})a_{-i} + (1-\beta_{-i})\bar{a} + e - q_i}{1+\delta_{-i}}$$

Because in equilibrium firm i knows its own output q_i and also a_i, β_{-i}, δ_{-i} and $\overline{a}$, the net information in the price signal is

$$(1-\alpha_{-i})a_{-i} + e. \tag{6}$$

Substituting for p into firm i's quadratic profit maximization problem yields, given firm $-i$'s conjectured linear supply schedule, and with firm i explicitly accounting for its impact on price,

$$\max_{q_i \in R} E\left[\left(\frac{a_i + (1-\alpha_{-i})a_{-i} + (1-\beta_{-i})\bar{a} - \delta_{-i}e - q_i}{1+\delta_{-i}}\right) q_i \middle| a_i, \overline{a}, p+e\right]. \tag{7}$$

Solving yields the conjectured linear form of output:

$$q_i = \frac{(1-\beta_{-i})\bar{a} + a_i + \lambda_i[(1-\alpha_{-i})a_{-i} + e]}{2} \tag{8}$$

where λ_i is the linear projection of $(1-\alpha_{-i})a_{-i} - \delta_{-i}e$ on the *net* information $(1-\alpha_{-i})a_{-i}+e$ in the noisy signal of price, $p+e$.

2.1. **Directly optimizing the supply functions.**

Because equilibrium strategies are linear, it is possible to simplify the analysis by assuming that strategies are linear and optimizing the supply functions themselves—i.e., optimize directly over the parameters α_i, β_i, and δ_i of the linear strategy rather than over output, as in (7)—to obtain the supply function. The first step in this approach is to posit that both firms have linear output functions as conjectured above:

$$\begin{aligned} q_i &= \alpha_i a_i + \beta_i \overline{a} + \delta_i(p+e) \\ q_{-i} &= \alpha_i a_{-i} + \beta_{-i}\overline{a} + \delta_{-i}(p+e) \end{aligned} \tag{9}$$

Thus, this approach takes as given that firm i's output is linear in its information. It is also key that the rival firm's *coefficients* α_{-i}, β_{-i}, and δ_{-i} of the rival firm's supply are taken as fixed by firm i, which is equivalent in this linear setting to taking the rival firm's supply *function* as fixed.

It further simplifies the analysis to posit that firms optimize over the intensity on net information in price, $(1-\alpha_{-i})a_{-i}+e$ directly via $\frac{\lambda_i}{2}$, rather than over δ_i. Substituting the conjectured linear structure into the price function, the objective becomes

(10)
$$\max_{\alpha_i,\beta_i,\lambda_i\in R} E\left[\left(\frac{(1-\alpha_i)a_i+(1-\alpha_{-i})a_{-i}+(1-\beta_{-i}-\beta_i)\overline{a}-\delta_{-i}e-\frac{\lambda_i}{2}((1-\alpha_{-i})a_{-i}+e)}{1+\delta_{-i}}\right)\right.$$
$$\left.\times\left(\alpha_i a_i+\beta_i\bar{a}+\frac{\lambda_i}{2}((1-\alpha_{-i})a_{-i}+e)\right)\Big|a_i,\overline{a},(1-\alpha_{-i})a_{-i}+e\right]$$

The next step is to carry through the expectation operator, exploiting the mutual independence of the shocks a_i, a_{-i}, $\overline{a}$ and e. This yields the modified objective

$$\max_{\alpha_i,\beta_i,\lambda_i\in R}\left\{(1-\alpha_i)\alpha_i\sigma_a^2+(1-\alpha_{-i})\left(1-\frac{\lambda_i}{2}\right)\frac{\lambda_i}{2}(1-\alpha_{-i})\sigma_a^2\right.$$
$$\left.+(1-\beta_{-i}-\beta_i)\,\beta_i\sigma_{\overline{a}}^2-\left(\delta_{-i}+\frac{\lambda_i}{2}\right)\frac{\lambda_i}{2}\sigma_e^2\right\} \tag{11}$$

Note that the $1+\delta_{-i}$ denominator has been dropped because it does not affect the optimisation for firm i. The first order conditions are

$$\begin{aligned}\alpha_i&=\frac{1}{2}\\ \beta_i&=\frac{1-\beta_{-i}}{2}\\ \lambda_i&=\frac{(1-\alpha_{-i})^2\sigma_a^2-\delta_{-i}\sigma_e^2}{(1-\alpha_{-i})^2\sigma_a^2+\sigma_e^2}\end{aligned} \tag{12}$$

Substituting these equations into the conjectured linear form of output duplicates the results obtained from output optimisation in (8).

2.2. **Equlibrium.** Substituting into the conjectured linear form for output in (9), in equilibrium the coefficients in (9) must be consistent with the coefficients in (12). Therefore, following the procedure in Bernhardt and Taub [13],

$$(a_i):\qquad \frac{1}{2}=\alpha_i+\delta_i\frac{(1-\lambda_{-i}(1-\alpha_i))}{2} \tag{13}$$

$$(a_{-i}):\qquad \lambda_i\frac{1-\alpha_{-i}}{2}=\delta_i\frac{(1-\lambda_i(1-\alpha_{-i}))}{2} \tag{14}$$

$$(\overline{a}):\qquad \frac{1-\beta_{-i}}{2}=\beta_i+\delta_i\frac{(\beta_i+\beta_{-i})}{2} \tag{15}$$

$$(e):\qquad \frac{\lambda_i}{2}=\delta_i\left(1-\frac{\lambda_i+\lambda_{-i}}{2}\right) \tag{16}$$

Imposing symmetry yields the solutions

$$\alpha=\frac{1-\lambda}{2-\lambda},\quad \beta=\frac{1-\lambda}{3-2\lambda},\quad\text{and}\quad \delta=\frac{\lambda}{2(1-\lambda)}. \tag{17}$$

as in Bernhardt and Taub [13]. In equilibrium, these weights must be consistent with the solution for the projection coefficient λ, which solves the nonlinear equation

$$\lambda = \frac{(1-\alpha)^2\sigma_a^2 - \delta\sigma_e^2}{(1-\alpha)^2\sigma_a^2 + \sigma_e^2}. \tag{18}$$

Substituting the solutions in (17) for the output weights into equation (18) yields a recursion in λ whose fixed point fully describes equilibrium outcomes:

Lemma 1. *A $\lambda \in [0,1]$ such that equations* (17) *and* (18) *are satisfied fully characterizes a symmetric linear equilibrium.*

Bernhardt and Taub [13] established existence by using a geometric approach to determine a fixed point of equation (18). To motivate how I establish existence in the dynamic game, I use a different approach here: I demonstrate that a contraction property holds for a variant of the recursion in (18). Net unrevealed private information—the forecast error variance on the net information in the price signal—is $(1-\lambda)^2$. Substituting for α and δ using (17) and defining $x \equiv 1-\lambda$, rewrite (18) as a recursion in x:

$$x = \left(\frac{1}{2}\frac{(1+x)^2\sigma_e^2}{(1+x)^2\sigma_e^2 + \sigma_a^2}(1+x)\right)^{1/2} \equiv T[x]. \tag{19}$$

Proposition 1. *$T(x)$ is a contraction mapping on $[0,1]$. Thus, a unique linear equilibrium λ exists.*

The proof is in Appendix B. The second-order condition for firm optimization holds if and only if $\lambda \in [0,1]$.[8]

2.3. **Signal jamming basic intuition.** In a Cournot duopoly with linear demand and with a full information demand shock $\overline{a}$, equilibrium output for firm i is

$$q_i = \frac{\overline{a}}{3} \tag{20}$$

Similarly, in the noisy price model If the noise is maximized then the price will have no useful information, and firms will behave as monopolists with respect to their privately observed shock, so that equilibrium output is the monopoly output

$$q_i = \frac{a_i}{2} \tag{21}$$

The intensities $\frac{1}{3}$ and $\frac{1}{2}$ thus provide benchmarks for assessing the noisy price equilibrium.

The noisy-price model generalizes these intensities. Recalling the optimal output function (8), the key observation to make concerning signal jamming is that in order for output on the public shock $\overline{a}$ to *increase* due to signal jamming, the intensity on β must *decrease* relative to its full-information value.

[8] A firm's second-order condition is $-1/(1+\delta)$ which becomes positive when $\lambda > 1$.

By solving for λ in terms of δ from (17) and substituting into (8), this yields the asymptotic behaviour

$$\beta \sim \frac{1+2\delta}{2(1+\delta)} \sim \begin{cases} \frac{1}{3} & \delta \to 0 \ (\sigma_e^2 \to \infty) \\ 0 & \delta \to \infty \ (\sigma_e^2 \to 0) \end{cases} \tag{22}$$

as the noise variance approaches infinity or zero respectively.

Thus, as the noise variance shrinks and the usefulness of the price signal increases, the firms *increase* the indirect intensity on the public shock in order to diminish price and thus to make it appear that demand has fallen; the *net* effect of the reduction in direct intensity and the increase in indirect intensity is to increase output—the intention of signal jamming.

2.4. **The equivalence of the intensity on learned information and signal jamming intensity.** Having established the basic mechanism of signal jamming, I now explore the interaction of signal jamming and learning. Again solving for λ in terms of δ from (17) and substituting into (8) , the total output intensity on the "learned" part of the a_i shock is

$$\frac{2\delta}{1+2\delta}(1-\alpha) \tag{23}$$

The "2" in the numerator of the second term comes from the inclusion of the rival firm's intensity on a_{it} as well as firm i's intensity.

Similarly for the public shocks, the total "signal jamming" intensity—that is, the indirect intensity via price—from both firms is

$$\frac{2\delta}{1+2\delta}(1-2\beta) \tag{24}$$

where the 2β term comes from the fact that both firms' signal jamming outputs are subtracted from the public demand shock. Thus, α and 2β are behaviourally identical—signal jamming takes place on learned information.

2.4.1. *Recursive feedback.* Returning to the equilibrium solution in (13), it is useful to begin with the β equation. The pair of equations can be written in recursive matrix form as

$$\begin{pmatrix} \beta_i \\ \beta_{-i} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} \beta_i \\ \beta_{-i} \end{pmatrix} + \begin{pmatrix} -\frac{\delta_i}{2} & -\frac{\delta_i}{2} \\ -\frac{\delta_{-i}}{2} & -\frac{\delta_{-i}}{2} \end{pmatrix} \begin{pmatrix} \beta_i \\ \beta_{-i} \end{pmatrix} \tag{25}$$

If δ_i is zero, the first part of this recursion is simply the standard Cournot reaction function equation with equilibrium $\beta = \frac{1}{3}$. The second part of the equation captures signal jamming: even though β is the intensity on $\overline{a}$, a positive δ—that is, if the price signal has useful information—then β responds. Notice that the response is strategic: a change in β feeds into δ_{-i} and then to β_{-i}. The result is that the output intensity on $\overline{a}$ is *reduced*—as discussed previously this then results in higher net output intensity, which is the expression of signal jamming.

It is also possible to develop some intuition from the equation for α_i, writing the equation from (13) recursively,

$$\alpha_i = \frac{1}{2} - \delta_i \frac{(1-\lambda_{-i}(1-\alpha_i))}{2} \tag{26}$$

If the price signal is extremely noisy, then δ_i will be near zero and the solution is simply the monopoly intensity, $\frac{1}{2}$. If the price signal is useful then an increase in intensity α_i *reduces* the signal observed by the rival via λ_{-i}, that is, the rival's *learning.* This encourages firm i to *reduce* its intensity α_i. But this is the same effect as with the intensity on the public demand shock: it is signal jamming. Except for the λ coefficient, which expresses the projection or learning about the firm's private shock by the rival—again, the learning—the structure is just as in the equation for β.

2.5. **Recursive learning and informational externalities.** Recalling the solution for λ_i from equation (12) and defining

$$\phi_i \equiv \frac{(1-\alpha_{-i})^2\sigma_a^2}{(1-\alpha_{-i})^2\sigma_a^2+\sigma_e^2} \tag{27}$$

we can write (12) as

$$\lambda_i = \phi_i - (1-\phi_i)\delta_{-i} \tag{28}$$

which is exactly equation (11) from [13]. Substituting from the equilibrium formula for δ_i in (13) yields

$$\lambda_i = \phi_i - (1-\phi_i)\frac{\lambda_{-i}}{2-\lambda_i-\lambda_{-i}} \tag{29}$$

This equation has two parts. The first part, ϕ, is the direct projection of excess demand from the unobservable shock on the noisy signal of that shock; this is conventional learning. The second part is more subtle. First, one can view $(1-\phi_i)$ as the estimate of the noise shock. The formula attempts to *subtract* this estimate from the learned information—but it is a subtraction of the other firm's learning.

The ratio $\frac{\lambda_{-i}}{2-\lambda_i-\lambda_{-i}}$ captures and expresses the idea of *feedback.* As the rival firm increases its output in response to the noisy signal in price (via λ_{-i}), this effectively *amplifies* the noise, and in response the quality of the signal in price is reduced, causing firm i to *reduce* its own intensity on its noisy signal. This effect is *further* amplified by the denominator term $-\lambda_{-i}$. This is how signal jamming interferes with learning.

The amplification from the denominator term expresses the common pool aspect of the noise. Because the λ_i and λ_{-i} terms are subtracted there is positive feedback from the intensity of both firms, which intensifies the response of both firms. Thus, if we think of extracting information—learning—from the noisy signal as a kind of fishing, both firms are reducing the quality of that signal from their intensities without internalising the cost to themselves.[9]

[9]That is, there is an externality connected with the aggregate information source—in this setting, the price—and that the firms, which because they want to collude, would like to internalise, but cannot. I note that this issue has been treated by Amador and Weill [5] in the context of economies where agents are small and non-strategic; Taub [57] treats a similar issue.

3. The dynamic model

In the benchmark stationary dynamic model, demand evolves stochastically according to first-order autoregressive processes. The underlying fundamental shocks $\{a_{it}, \overline{a}_t, e_t\}$, $t \in (\dots, -1, 0, 1, \dots)$, $i \in \{1, 2\}$, are serially-uncorrelated, zero-mean Gaussian processes, with variances σ_a^2, $\sigma_{\overline{a}}^2$ and σ_e^2, respectively. Extending the static model, the period-t price function becomes

$$\pi(q_{1t}, q_{2t}, X_t) = A_1(L)a_{1t} + A_2(L)a_{2t} + B(L)\overline{a}_t - (q_{1t} + q_{2t}), \tag{30}$$

where $X_t \equiv \{A_1(L)a_{1t}, A_2(L)a_{2t}, B(L)\overline{a}_t, e_t\}$ is the driving process vector, and q_{it} is period-t output by firm i. Here, L denotes the lag operator, i.e., $Lx_t = x_{t-1}$, and, for example, $A_i(L)$ is a linear function of the lag operator, $A_i(L)a_{it} = \sum_{j=0}^{\infty} \rho^j a_{i,t-j}$, so that $A_i(L) = \frac{1}{1-\rho L}$. Analogously, $B(L)\overline{a}_t$, with $B(L) = \frac{1}{1-bL}$, is a publicly-observed common-value demand process; $A_i(L)a_{it}$, with $A_i(L) = \frac{1}{1-\rho L}$, is a common-value demand-shock process privately observed by firm i; and e_t is a Gaussian i.i.d. unobservable noise shock that is also independent from the demand shocks. Firm i does not see its rival's demand process $A_{-i}(L)a_{-it}$.

Firms have discount factor η, which is assumed to satisfy $|b| < \eta^{-1/2}$, $|\rho| < \eta^{-1/2}$, and $|\phi| < \eta^{-1/2}$ so that expected discounted sums such as $E[\sum_{s=0}^{\infty} \eta^s (A_i(L)a_{i,t+s})^2]$ converge. I assume that there are no costs of production in order to reduce the complexity of the model.

As in the static model, firms compete in supply schedules. Denote the realized period-t price that firms receive for their output by p_t, the private history of the shocks observed by firm i by $X_i^t \equiv \{a_{i,t-s}, \bar{a}_{t-s}\}_{s=0}^{\infty}$, and the history of its current and past outputs by $q_i^t = \{q_{i,t-s})_{s=0}^{\infty}$. Denote the price history by $p^t \equiv \{p_{t-s}^P\}_{s=0}^{\infty}$.

A supply schedule for firm i is a differentiable function $Q_t^i : (p^t + e^t; X_i^t, q_i^{t-1}) \mapsto \mathbb{R}$ that maps each period t the price signal, and histories of shocks observed by firm i, prices and past outputs into an output level. A price function $\pi : (q_1, q_2, X_t) \mapsto \mathbb{R}$ is *market clearing* if it is consistent with the supply schedules:

$$q_{1t} = Q_t^1(p^t + e^t; X_1^t, q_1^{t-1}), \quad q_{2t} = Q_t^2(p^t + e^t; X_2^t, q_2^{t-1}), \quad \text{and} \quad p_t = \pi(q_{1t}, q_{2t}, X_t) \tag{31}$$

This implicitly defines a fixed point problem in the space of functions containing π, Q^1, and Q^2.

As with the static model I assume that current and past realized profit do not result in any improvement in the signal of price.

I only characterize stationary equilibrium path outcomes. With Gaussian shocks, all possible price histories are consistent with some equilibrium path because the Gaussian shocks have support over the entire real line, so there are no off-equilibrium beliefs to specify. [10]

I solve for an equilibrium in which the supply functions are linear and stationary, taking the form

$$Q^i(p^t + e^t, X_i^t) = \alpha_i(L)A_i(L)a_{it} + \beta_i(L)B(L)\overline{a}_t + \delta_i(L)(p_t + e_t) \tag{32}$$

[10]For a related discussion see Foster and Viswanathan [22], p. 1446.

Definition 2. *A stationary linear equilibrium is a pair of supply functions $Q^{i*}(\cdot;\cdot)$ with linear weighting functions $(\alpha_i(L), \beta_i(L), \delta_i(L))$ satisfying* (32) *for $i \in \{1, 2\}$, and a market-clearing price function $\pi^*(\cdot, \cdot, \cdot)$ such that for each current price p_t and history (X^t, X_1^t, X_2^t) and price signal history $p^t + e^t$: $Q^i(p^t + e^t; X_i^t)$ maximizes firm i's expected profit given p^t, and X_i^t; and prices and output are market clearing, satisfying* (31) *for all $t = \dots, -1, 0, 1, \dots$.*

Solution procedure. To find the linear equilibrium I first follow in parallel the steps used in the static model: conjecture that firm i's rival's output is a linear function of its information history and substitute the rival's posited linear output functions into the price function, which is therefore also linear in the history of the fundamental processes. Firm i's optimization problem inherits the linearity of the price function, preserving the linear-quadratic structure of its objective. I then show that firm i's best response is linear in its information history.

I then replicate the strategy in the second approach to the firm's problem in the static model: optimize over the supply function itself,[11] which in a linear setting translates to optimizing the intensities, which are now functions of the histories of observed public and private shock realizations and of the history of the price signal.[12] I then translate the optimisation problem to the frequency domain and solve.

To begin I verify the linearity of the output functions.

Lemma 2. *Let (i) firm $-i$'s supply function $Q^{-i}(\cdot;\cdot)$ be a linear and stationary functional of the history X_{-i}^t, the history of prices p^t, and (ii) let firm i's best response in future periods $t+1, t+2, \dots$ be a stationary linear functional of its information history X_i^{t+s}, $s = 0, 1, \dots$ Then firm i's optimal output is a stationary linear functional of X_i^t and p^t.*

The first-order condition for the time-domain objective is difficult to solve due to the infinitely many future values $q_{i,t+s}$ appearing in the first-order condition, interacting quadratically with terms at other lags. To proceed, I exploit the equivalence of firms' optimization over the functions in the time domain and their optimization in the frequency domain when the optimal supply functions are linear. In the frequency domain these functions are called *filters*, and can be manipulated as algebraic objects.[13] The conditional expected profit objective in the time domain maps into an inner product that is a function of these filters. A firm's expected profit maximization problem is then a variational problem that is solved by the optimal filter, where firms compete in these filters directly.

Transforming the objective to the frequency domain. To transform the model to the frequency domain, first apply Lemma 2 to firm i, that is, conjecture that the rival's output intensity process is determined by linear functions of the lag operator α_{-i}, β_{-i} and δ_{-i}.

[11] I establish that this is equivalent to conventional time-domain optimization in Appendix A.

[12] Because the functions that are being chosen explicitly act on information processes, including endogenous signals, beliefs are automatically taken into account. Thus, any equilibrium that is determined by the fixed point argument is automatically sequentially rational.

[13] This algebraic character of the frequency domain is analogous to the algebraic character of the Laplace transform methods used to solve differential equations. In control systems engineering the frequency-domain functions would be called transfer functions, with the term filter reserved for the physical implementation of the solution.

Analogously with the procedure in the static model, substitute the linear output function of firm i as a function of its information and the conjectured linear form of its rival (see Appendix C, equation equation (72)) into the price function to obtain its linear structure

(33)
$$p_t = \pi(q_{1t}, q_{2t}; X_t) = (1 + \delta_1(L) + \delta_2(L))^{-1}\Big((1 - \alpha_1(L))A_1(L)a_{1t} + (1 - \alpha_2(L))A_2(L)a_{2t} + (1 - \beta_1(L) - \beta_2(L))B(L)\overline{a}_t + e_t\Big).$$

Next, substitute both the conjectured output strategy for firm $-i$ and firm i's best response into price. Substituting this solution for price into firm i's output function yields its output as a linear function of the history,

$$(34)\quad q_{it} = \alpha_i(L)A_i(L)a_{it} + \beta_i(L)B(L)\overline{a}_t + \delta_i(L)(1 + \delta_1(L) + \delta_2(L))^{-1}\Big((1 - \alpha_1(L))A_1(L)a_{1t} + (1 - \alpha_2(L))A_2(L)a_{2t} + (1 - \beta_1(L) - \beta_2(L))B(L)\overline{a}_t + e_t\Big).$$

I next transform the objective to the frequency domain using these linear expressions.

Details of the mapping. The firm's profit is the sum of the discounted expected profit; in each period the discounted time-t expected profit term $E[p_t q_{it}]$ appears (omitting conditioning and discounting). Each of the terms p_t and q_{it} is the sum of functions operating on the fundamentals a_{1t}, a_{2t}, $\overline{a}_t$, e_t and so on, using equations (33) and (34). For example, p_t contains the term

$$(35)\qquad (1 + \delta_1(L) + \delta_2(L))^{-1}(1 - \alpha_i(L))A_i(L)a_{it},$$

operating on a_{it} in equation (33), which is cross-multiplied by

$$(36)\qquad \alpha_i(L)A_i(L)a_{it} + \delta_i(L)(1 + \delta_1(L) + \delta_2(L))^{-1}(1 - \alpha_i(L))A_i(L)a_{it}$$

from firm i's output q_{1t} in equation (34), also operating on a_{it}. The cross-products of these elements with all other terms are zero because the underlying stochastic processes are uncorrelated. After carrying out the complicated multiplications of the functions $\alpha_i(L)$, $A_i(L)$ and $\delta_i(L)$, there will be a summation of terms that can be abstractly represented as

$$(37)\qquad E\left[\left(H_j L^j a_{i,t+s}\right)\left(G_k L^k a_{i,t+\tau}\right)\Big|(a_{it}, \overline{a}_t, p_t + e_t)\right]$$

Applying the lag operator and bringing the expectation operator inside then yields

$$(38)\qquad H_j G_k E\left[a_{i,t+s-j}a_{i,t+\tau-k}\Big|(a_{it}, \overline{a}_t, p_t + e_t)\right]$$

where H_j and G_k abstractly represent the complicated products of the coefficient terms in the functions $\alpha_i(L)$, $A_i(L)$ and $\delta_i(L)$.

As long as $s-j=\tau-k>0$ this reduces to H_jG_k, otherwise it is zero, as the fundamental innovations a_{it} are i.i.d.. Importantly, the conditioning does not add complications because the innovation processes, which, again, are i.i.d., cannot be predicted from the current and past values of the realised shocks and noisy signals.

The remaining task is to determine whether there is a usable structure in the remaining terms of the objective. To generate this structure the following equivalence holds:

$$E\big[a_{i,t+s-j}a_{i,t+\tau-k}\big] \sim \sigma_a^2 \frac{1}{2\pi i}\oint z^s z^{-\tau}\frac{dz}{z} = \begin{cases} 0, & s-j\neq \tau-k \\ \sigma_a^2, & s-j=\tau-k, \end{cases} \tag{39}$$

where the integral is a contour integral around the unit circle in the complex plane. The intuition of the contour integral and why the equivalence holds is presented in Appendix A, but the main conclusion is that the conventional time-domain objective—the conditional expectation of a complicated summation of future quadratic terms—is exactly equivalent to a contour integral. Furthermore, the contour integral is of a specific type: it is a *convolution*, which defines an inner product, involving the functions comprising the firms' supply functions and the price process.

These two terms in (35) and (36) thus interact as an inner product, appearing as the convolution integral

$$\frac{1}{2\pi i}\oint D(1-\alpha_i)(\alpha_i^* + \delta_i^* D^*(1-\alpha_i^*))A_iA_i^*\sigma_a^2\frac{dz}{z}, \tag{40}$$

using the definition

$$D(z)\equiv (1+\delta_1(z)+\delta_2(z))^{-1}. \tag{41}$$

and where the "*" notation denotes the conjugate function, which has negative powers of z,

$$D^*\equiv D(\eta z^{-1}) \tag{42}$$

and so on for the other functions.[14]

Because the fundamental innovations a_{it}, a_{-it}, $\overline{a}_t$, and e_t are uncorrelated, the frequency domain formulation of the objective cleaves into parts attached to the variance of each of the innovation processes a_{it}, $\overline{a}_t$ and e_t. Following the same procedure used to obtain (40), one obtains the frequency domain version of firm i's objective:

(43)

$$\max_{\alpha_i,\beta_i,\delta_i}\frac{1}{2\pi i}\oint\Bigg(D(1-\alpha_i)(\alpha_i^*+\delta_i^*D^*(1-\alpha_i^*))A_iA_i^*\sigma_a^2 + D(1-\alpha_{-i})\delta_i^*D^*(1-\alpha_{-i}^*)A_{-i}A_{-i}^*\sigma_a^2$$
$$+D(1-\beta_i-\beta_{-i})\left(\beta_i^*+\delta_i^*D^*(1-\beta_i^*-\beta_{-i}^*)\right)BB^*\sigma_{\overline{a}}^2 + (D-1)\delta_i^*D^*\sigma_e^2\Bigg)\frac{dz}{z},$$

and symmetrically for firm $-i$. See Appendix C for the details of this derivation.

[14]The coefficients in D^* are also the complex conjugates of the coefficients in $D(z)$, however due to the factorization property discussed later it is not necessary to highlight this fact.

The optimization in (43) is over the intensity *functions* α_i, β_i, and δ_i, taking as given the rival's intensity functions α_{-i}, β_{-i}, and δ_{-i}. Just as in the static model, where the choice of the intensities α, β and δ was equivalent to optimising over output, this approach is equivalent to the time-domain approach; the general equivalence of frequency-domain and time-domain optimisation is established and discussed in more detail in Appendix A.4.[15]

Variational derivatives. I set out the detailed variational derivatives—the Euler equations—of the frequency domain objective (43) in Appendix C, equation (79). These equations are asymmetric, reflecting the ability of firms to weight histories of signals, but not future realizations of the signals; Euler equations of this type are called Wiener-Hopf equations. This asymmetry was earlier expressed in the first-order condition (78), and the solution strategy must account for the asymmetry. That is, the variational first-order condition nominally resembles the first-order condition for a static quadratic optimization problem, which might be abstractly represented as an equation

$$My = Bx,$$

where the objective is to solve for y. This would be conventionally done by inverting the M matrix. However, in the frequency domain, this inversion cannot be done because it implicitly requires putting weights on future realizations of the history, which are inherently unobservable.

To circumvent this inversion problem, one follows four steps: (i) factor the M matrix (keeping in mind that the M matrix is a matrix of *functions*) into the product $F'F$ of two matrices, F and F', where F corresponds to the weighting of histories, and F' corresponds to the (infeasible) weighting of *future* realizations; (ii) invert F'; (iii) apply a projection to the resulting right-hand side—the $[\cdot]_+$ operator that eliminates terms that weight future histories;[16] and finally (iv) invert F; importantly, the inverse of F only weights *past* and *present* but not future realizations. The resulting formula is equivalent to constructing a linear least squares projection—a regression—on the history.

It is an important detail that the factorisation step, step (i) above, is guaranteed to have a solution F that is (a) analytic on the domain of interest, (b) is also invertible on that domain, and (c) has real coefficients. Thus, the inversion that takes place in steps (ii) and (iv) can always be carried out. This result is due to a theorem of Rozanov [46].

---

[15] The frequency-domain approach limits the controls to *stationary* linear strategies, in the sense that the same choice of the linear filters α_i, β_i, and δ_i is applied in each period when mapped into the time domain, representing a fixed point of the time-domain first-order condition (78). Thus, if there are also "bubble" solutions for the output process, i.e., equilibria in which the α_i, β_i, and δ_i functions are linear but time varying, the approach will not find them. It is also implicit that the solutions are dynamically consistent, that is, the frequency-domain solution finds the linear filter that would be replicated in every period, conditional on its future structure, in a time-domain approach; this is a quotidian result for additively separable systems like the one here.

[16] The projection or "annihilator" operator, $[\cdot]_+$, eliminates terms with negative powers of z from the Laurent expansion of a function: if $f(z) = \cdots + b_{-2}z^{-2} + b_{-1}z^{-1} + b_0 + b_1z^1 + b_2z^2 + \ldots$, then $[f]_+ = b_0 + b_1z^1 + b_2z^2 + \ldots$. The annihilator operator accounts for the fact that firms can weight histories of observed signals in their strategies, but not the yet-to-be-observed future realizations of signals.

Solution of a firm's variational optimization problem and equilibrium. It is useful to gather terms by implicitly defining two objects, F and J, which solve

$$F^* F \equiv D(1-\delta_1^* D^*) + D^*(1-\delta_1 D) \tag{44}$$

$$J^* J \equiv (1-\alpha_2^*)(1-\alpha_2) A_2 A_2^* \sigma_a^2 + \sigma_e^2. \tag{45}$$

which have, as discussed above, factorizations F and J. The static analogue of F is $2(1-\lambda)-\lambda$: it is the projection coefficient structure corresponding to the net information in the noisy price signal (the $2(1-\lambda)$ term) after a firm has extracted information from the price signal (the $-\lambda$ term). The function J is the filter characterizing the information process from firm 1's observation of the price signals; it is the dynamic analogue of the net information $(1-\alpha_{-i})a_{-i}+e$ in the price in the static setting.

Taking the variational derivatives and exploiting symmetry to solve for the Wiener-Hopf equations, yields the optimal filters:

Proposition 2. *In a symmetric equilibrium firm i's filters on its direct information sources, a_i and $\bar{a}$ are given by*

$$\begin{aligned} \alpha &= 1 - F^{-1}A^{-1}\left[F^{*-1}D^* A\right]_+ \\ 1-\beta &= \frac{1}{2} + \frac{1}{2}F^{-1}B^{-1}\left[F^{*-1}(1-\beta)D^* B\right]_+ \end{aligned} \tag{46}$$

Output weights on price signals satisfy the recursive system

$$J^* J = F^{-1}\left[F^{*-1}D^* A\right]_+ \left[F^{*-1}D^* A\right]_+^* F^{*-1} + \sigma_e^2 \tag{47}$$

$$D = \frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+ + \frac{1}{2}J^{-1}\left[J\frac{D+D^*}{1+D^*}\right]_+ . \tag{48}$$

Lemma 3. *If analytic functions $\{\alpha,\beta,\delta\}$ in $H^2[\eta]$ satisfy* (46), (41), (47), (48), *then the time-domain version of $Q^{i,t}$ defined in* (32) *is a stationary linear equilibrium.*

Proof. The result is immediate using the equivalence of frequency domain optimization with time-domain optimization, as consistency is also satisfied. □

Proposition 3. *A stationary linear equilibrium exists.*

The proof is in Appendix D. The fixed point argument uses the recursive system (47)–(48). Recalling the relationships between the static weights and λ in equation (17) of the static model, D as defined in (41) is the dynamic analogue of $1-\lambda$, suggesting that the recursion in (48) is analogous to its static counterpart. Denoting the right-hand side of (48) by $S(D)$, write the recursion as

$$D = S(D),$$

defining a recursion in D (equation (47) is ancillary). I show that $S(D)$, which is a continuous mapping, is bounded by a function $T(D)$; and that this bounding function $T(D)$ is itself a contraction on the unit disk and as such has a unique fixed point. It follows that S has a fixed point. I then use the Szegö form of the function to establish that the fixed point is not at $D = 0$. I believe this approach to demonstrating the existence of a fixed point to be original.

4. Characterization of dynamic signal jamming and learning

In this section I characterise the behaviour of the dynamic model. There are some initial building blocks. As a first step I show that even though the firms engage in signal jamming on public information, neither the public information fundamental shocks nor the outputs driven by those shocks affect the behaviour of the firms toward their privately observed shocks. The converse is not true however: signal jamming on public information is fully shaped by the nature of the private shocks.

As a related point I demonstrate that the intensities α and β are not scalars, that is, the equilibrium intensity filters alter the underlying autoregressive structure of the fundamental demand processes to determine output.

With these results in hand I can then characterise the interrelationships between learning and signal jamming. Not surprisingly the properties of static signal jamming are retained, however learning is now identified with low frequencies. At low frequencies, where learning is expressed, the output intensity on public and private learned information converge. Moreover, one can decompose output into direct, indirect and total output, and this can be captured using frequency domain methods: surprisingly, as the rival learns about private shock realisations, it steps up its intensity on them, treating them like public shocks, whilst the firm actually able to observe the shocks *steps back* its intensity.

The public component of the demand process, $B(L)\overline{a}_t$, can be netted out of price directly, leaving the net information in price independent of $B(L)\overline{a}_t$. It follows that $B(L)\overline{a}_t$, does not affect optimal output weights on prices, and thus does not enter the equilibrium functions F, J or D. The following is immediate.

Corollary 1. *The public information component of demand affects neither the filtering of the price process by firms, nor firm output weights on private information.*

The proofs of this and other propositions from this section are in Appendix E.

The converse to Corollary 1 is not true—the private information components of demand affect output weights on publicly-known demand both directly, and indirectly via the output weights on privately-observed demand, via the functions D and F that appear in the solution for β—this is signal jamming. Corollary 1 also implies that one could have added any deterministic component to demand, and solved for the equilibrium: this deterministic component would have no effects on the portions of output that reflect private information or information contained in prices.

I next establish that the equilibrium direct intensity filters are not scalar constants, and have high-order autoregressive structure, with smaller autoregressive parameters than those of the fundamental demand shock processes, indicating that the autoregressive structure of output is fundamentally altered.

Proposition 4. *The equilibrium output intensity filters α_i and β_i are not scalar-valued: output intensities are not just amplifications of the dynamic shock processes. This autoregressive structure reflects strategic behavior and is not just the result of signal extraction alone.*

The non-scalar nature of output responses is not just due to the fact that the firms filter the price signals via signal extraction, i.e., construct estimates of the exogenous driving processes, as they would in a competitive noisy rational expectations equilibrium economy with informationally-small firms. Firms also internalize the fact that they are informationally large, actively signal *jamming* to influence a rival's inferences. This alters the time series structure of output in ways that signal *extraction* alone does not. In particular, I prove that were firms solely engaging in signal extraction from prices, then the autoregressive coefficients of the output processes would equal those of the exogenous demand processes. I then prove that the autoregressive coefficients of these two processes differ and that any equilibrium output process is of infinite autoregressive order.

4.1. **Theoretical results on dynamic signal jamming.** I now carry out a couple of thought experiments in which I vary the dynamic character of the fundamental shock processes, specifically the public shock process $B(L)\overline{a}_t$ and the privately observable shocks $A_i(L)a_{it}$.

In the first experiment, posit that the privately observable shocks are i.i.d, that is, that the function $A(z)$ is a scalar constant. From Corollary 1, the public information process does not affect the output intensity on private information, and therefore it is possible to characterise the output intensity on the private shock processes using only the characteristics of the private side. Applying the "annihilator" lemma, Lemma 6 from Appendix A, it is immediate that the private output intensity $\alpha(L)$ will be a scalar constant as well, and also that the function J will be a scalar constant. (See equations (46) and (47).) This in turn implies that δ and D will be scalar constants, and then finally that F is a scalar constant. (See equations (44) and (48).) Examining the formula for β in Proposition 2, we see that all terms on the right hand side are then scalar constants with the exception of the filter $B(z)$. As a result, the annihilator formula becomes the identity, and we have

$$\begin{aligned}\beta &= F^{-1}B^{-1}\left[F^{*-1}(1-\beta)\left(D(1-\delta^*D^*)-D^*\delta D\right)B\right]_+ \\ &= F^{-1}B^{-1}F^{-1}(1-\beta)\left(D(1-\delta D)-D\delta D\right)B \\ &= F^{-1}F^{-1}(1-\beta)\left(D(1-\delta D)-D\delta D\right)\end{aligned} \tag{49}$$

which is a scalar constant. Thus, regardless of the dynamic structure of the publicly observable process, the intensity on the process is a scalar constant. The intensity on the public shocks reflects conventional static reasoning: firms react only to the *current* realization of the public shock to determine output on that shock, even if the shock is serially correlated.

For the second experiment reverse the situation: let the privately observable shocks be serially correlated, for example $A(z) = \frac{1}{1-az}$, but the publicly observable shocks are i.i.d., that is, $B(z)$ is a scalar constant. In that case the endogenous functions α, J, δ, D and F will all have a nontrivial structure; from other theorems that they have infinitely many poles for example. On the other hand the intensity on the public process, β, would, if there were no further influences, remain a scalar constant. That scalar constant might exceed the full-information value of $\frac{1}{3}$ due to signal jamming, however this is not a solution. Examining the formula for β in Proposition 2, we see that the functions δ, D and F enter the formula, and they do not cancel as they are not scalars. We can assert that β must therefore have at least one pole determined by the δ process. Thus, we have the following proposition:

Proposition 5. *Let the public demand shock process $B(L)\overline{a}_t$ be i.i.d., that is, $B(L)$ is a scalar constant, and let the private demand shock process $A(L)a_{i_t}$ be serially correlated. Then direct output on the public shock $\beta(L)B(L)\overline{a}_t$ will be serially correlated.*

This is dynamic signal jamming: even though the underlying public demand shock is i.i.d., the dynamic structure of the *output* on the public shock is driven by the structure of the private shock process. The reason for this is that each firm wants to *disguise* its output on its private process by making the output process on the public shock indistinguishable from the output on the private process.[17]

4.2. **Signal-jamming in the dynamic model.** Recall from the development of the static model that firms carry out signal jamming on the public signal by *reducing* the intensity β. The β formula is framed differently in the dynamic model, but a crude argument establishes that similar behaviour of β occurs in the dynamic model. From Proposition 2 the formula for the equilibrium β is (under symmetry) can also be written as

$$\beta = F^{-1}B^{-1}\left[F^{*-1}(1-\beta)(F^*F-D^*)B\right]_+ \tag{50}$$

Proceeding informally, now consider this equation whilst ignoring the annihilator operator, allowing cancellations which then yield

$$\beta = (1-\beta)(1-F^{-1}F^{*-1}D^*) \tag{51}$$

Similarly in informal fashion, ignoring the star-conjugation, we have

$$F^*F \sim 2D(1-\delta D)$$

so

$$F^{-1}F^{*-1}D^*\ \frac{D}{2D(1-\delta D)} \sim \frac{1+2\delta}{2(1+\delta)}$$

duplicating equation (22) from the static model, and yielding similar asymptotic behaviour.

Applying this intuitive result in the dynamic model we are interested in discovering the extent to which signal jamming *reduces* the intensity β, and at which frequencies. As will be demonstrated in the numerical simulations this reduction does take place, at low frequencies.

There is an additional effect in the dynamic model as well: as firms' estimates of their rival's private demand shock innovations improve at long lags, that is as they learn, they treat the learned information as if it is public information. Following the signal jamming logic above, they then relatively *reduce* the intensity on this learned information. The

[17]This echoes the inconspicuousness findings of the literature on informed trading on private information in stock markets as in the model of Kyle [37].

treatment is not necessarily identical however because the dynamic structure of the learned information might differ from the public shocks, that is, the serial correlation might differ.

However it is essential to keep the distinction between the *direct* intensity on the public demand shock, namely β, and the *total* intensity on that demand shock. As a result of signal jamming the *total* intensity on the publicly observable process *increases* at low frequencies, and similarly, because of the learning that takes place by the rival, so does the intensity on the privately observed shock. Thus the persistence of output *increases* overall relative to its static Cournot structure.

5. Numerical examples

In this section I explore three basic numerical examples. In the first example I suppose that the noise variance is so high that the signal from price is very noisy, leaving little scope for signal jamming and learning. In the second and third examples I reduce the noise so as to make the price signal usable and then explore the how learning occurs and the consequence of signal jamming on the dynamics of output. In the first of these latter two examples I suppose that the privately observed demand shocks are highly persistent whilst the publicly observable shock is not; in the final example I reverse the situation.

I characterise the results in two ways. The first way is to analyse the endogenous poles that emerge from the numerical estimate of the equilibrium output intensity filters. The key properties of the model are then determined by the dominant poles.

The second method for characterising the result is to examine the spectral densities of the relevant filters. I compute and characterise the spectral densities for both the intensity filters and the output processes that result from applying these filters to the fundamental shock process filters A_i, B, and the noise process. (More details on the interpretation and algebraic properties of spectral densities is provided in Appendix G.)

In order to numerically estimate the model I use so-called state-space methods, adapted from engineering control theory to simulate and iterate the recursion in equation (48). These methods suppose that the stochastic processes have an autoregressive-moving average structure, but can otherwise be arbitrary vector processes, i.e., processes that can be represented as

$$x_t = Ax_{t-1} + Bu_t, \tag{52}$$

where x_t and u_t can be vector processes, and A and B are appropriately conformable matrices. In engineering settings, x_t is the state process, and the u_t process is a serially uncorrelated and i.i.d. process, i.e., white noise. When x_t and u_t are scalar-valued and A and B are scalar constants, this is just an AR(1) process.

To analyze the dynamics of output, one could find the fixed point of the recursion in (48), use this to calculate the functions α, β, and γ, and then use those formulas to calculate the equilibrium weights on the input processes. These properties would be embodied in the poles—the eigenvalues of the A-matrix—of the state space versions of the functions. The proof of Proposition 4 reveals that there are infinitely many such poles, and to establish the pattern of the eigenvalues of the equilibrium A-matrix for δ, one must establish how the eigenvalues are affected by the recursion in (48).

To numerically approximate the equilibrium requires imposing an algorithm that trims quantitatively unimportant terms after each iteration of (48). I use methods developed in the engineering literature for such approximations. This literature also establishes error bounds for the approximations. These methods are described in greater detail in Appendix F.

5.1. **Maximum noise.** As the first example I examine the case in which the noise process has an extremely large variance relative to the variances of the fundamentals of the demand shock processes. When noise is large, there will be no useful information in price ($\delta(z) = 0$), so there will be no signal jamming. As a result the firms will treat their private shocks as monopolists, choosing output intensity filter as simply the scalar $\alpha(z) = \frac{1}{2}$, so that the output process filter $\alpha(z)A(z) = .5A(z)$. Because there will be no signal jamming, the output intensity on the public shocks will be the scalar from the static duopoly equilibrium, $\beta(z) = \frac{1}{3}$, and the resulting direct output process will be $\beta(z)B(z) = \frac{1}{3}B(z)$ from each firm.

I confirm these assertions in the simulations. To begin I choose tolerances for the state space operations in the simulations, presented in Table 1.

The numerically calculated output intensity functions are presented in Table 3. With high noise, there is essentially no signal jamming: the intensity on A is close to the static monopoly value of .5 with very little serial correlation, and the intensity on B is close to the static full-information duopoly value of .33, again with very little serial correlation.

These effects also show up in the spectral densities. Examining the spectral densities related to the private process $A(L)a_{it}$ in Figure 8 it is evident that there are essentially no indirect effects because the signal in price is so weak. The spectral density for total output thus closely mimics the spectral density for the monopoly output process $\frac{1}{2}A$.

The results are similar for the public process, however it is evident in Figure 9 that there is a noticeable signal-jamming effect: the B process is slightly amplified at low frequencies, relative to the static full-information intensity of .33.

Again because there is very little usable information in the price signal, Figure 10 shows that the output on noise is basically zero.

Finally, Figure 11 shows that output tracks the combined full-information output from the two demand shock processes.

5.2. **Persistent a_t process, low-persistence $\overline{a}_t$ process.** For the second example I examine the result of Proposition 5 numerically using the parameterization in Table 4.

Thus, the privately-observed demand shock processes $A(L)a_{it}$ are moderately positively serially correlated, while the publicly observable demand shocks $B(L)\overline{a}_t$ are much less serially correlated. (The serial correlation is not reduced to zero to avoid numerical instabilities.) The noise is also set at a moderate level—small enough so that there is some incentive for signal jamming.

The main prediction of Proposition 5 is that the firms will produce output on the publicly observable shocks that is much more serially correlated than the shock process itself, and that the pattern of output will resemble output on the private shocks. This is because if there is a difference in the serial correlation pattern of output on the two shocks separately, then signal extraction can extract the private shock more effectively. These effects are evident in the filters for output on the public shock in Table 5. The private direct intensity filter $\alpha(z)$ has main pole 2.19, which exceeds 2.0, the pole of the fundamental; the direct intensity for the public shock also has a nontrivial term with the same pole. Thus, the output on the public shock is persistent even though the underlying shock is not persistent.

The direct intensities are ultimately of less interest than the total output that combines the direct intensities with the indirect intensities from price, including the rival's output on a firm's private shocks operating through price. There we see a direct component that is driven by the same pole, 2.19.

It is also evident that there is significant serially correlated output on the noise alone, which is assumed to be serially uncorrelated.

It is also possible to draw conclusions by comparing the spectral densities of the processes in the model. Figure 12 displays the results related to the private shock process $A(L)a_{it}$. The first panel displays the spectral density of the private shock process and also the direct intensity filter for that shock, $\alpha(z)$. One can see that there is positive serial correlation in the shock process because of the hump in the middle, reflecting higher power at low frequencies. The spectral density of the direct intensity filter $\alpha(z)$, by contrast, de-emphasizes the low frequencies, as is evident from the dip in the middle of the plot. Again, because learning takes time, the learned information is signal jammed at low frequencies; thus, the de-emphasis at low frequencies reflects the reduction in intensity due to signal jamming on learned information.

The second panel compares the spectral density of the output resulting from the application of the filter to the shock, and compares it to the spectral density of the output process that would result from the firm choosing monopoly output on the shock, which is simply to multiply the shock in each period by $\frac{1}{2}$. It is evident that the low frequencies have been reduced relative to the high frequencies, as is expected from the shape of the spectral density for α. Thus, the filter can be seen as accelerating output and rendering it less serially correlated than the input process.

It is also important to notice that the variance of the output resulting from the direct intensity—corresponding to the area under the spectral density—is reduced relative to the input process. As we will see this attenuation of output is picked up in the other dimensions of the model.

Indirect effects. In the second row of Figure 12 we see the indirect intensity on the private process resulting from the rival's output on the noisy price signal. In contrast to the direct intensity, the indirect intensity emphasizes low frequencies, which relatively increases the long run output stemming from the private shock process. As a result, the spectral density of the output process driven indirectly via price (second panel in the second row of the figure) has relatively high persistence.

In the third row of Figure 12 we see the firm's total "own" intensity and output, that is, the output resulting from the application of the direct intensity and also the output from the firm's own indirect output from price. The second panel shows clearly that the "own" output is significantly attenuated at low frequencies, in contrast to the indirect output effect in the second row. This emphasizes that the firm "steps back" its own output in response to the rival's output on the signals in price, at low frequencies. At the lowest frequencies the rival completely takes over the output on the private shock!

Total output. The fourth row displays the spectral densities for the total intensities and the total output on the private shocks. The net effect is to reduce the overall volatility of the output on private shocks, but to significantly increase its persistence relative to the underlying demand shock. Thus, even though there is acceleration from the direct intensity, the net effect is deceleration, that is, increased persistence of output relative to the demand shock process.

Output on public demand shocks—signal jamming. Figure 13 displays the spectral densities associated with the publicly observable demand shock process $B(L)\overline{a}_t$. In the first panel of the first row we have the underlying process as it would appear in a full-information model: the Cournot intensity is $\frac{1}{3}$. As with the private shocks, the direct intensity attenuates low frequencies; the right panel shows that the output process has low frequencies attenuated and high frequencies intensified, having the overall effect of *reducing* the serial correlation of output. It is important to note that this does not contradict the content of Proposition 5, because the proposition refers to the direct intensity, whilst the total output includes indirect effects as well.

Indirect effects. As with the private shocks, the indirect effect of the price signal is to accentuate the low frequencies.

Total output. The right panel in the second row displays the spectral density for the total output on the public shock process. In the total output effect the low frequencies are blocked; because the spectral density has an inverted shape the output process has negative serial correlation. The indirect effects are swamped by the direct intensity, which has the negative serial correlation.

As with the private shocks, the total volatility—the area under the spectral density—is attenuated relative to the volatility that would prevail in a full-information Cournot setting.

Output from noise. Finally, consider output from noise in Figure 14. It is evident that the output from the noise process will be highly persistent, even though the noise process itself is i.i.d.. Moreover, there is a significant amount of power in the output on noise; compared to the output on the public shocks the power is very high. Thus, much of the signal jamming effects show up in the noise process rather than the public process per se.

Total output. Combining these effects in Figure 15, the overall power of the output spectrum is reduced relative to the full-information outputs that would result from the monopoly and Cournot output processes; this is because the noise causes an overall pulling back of output in general. In addition to this however we see that persistence is relatively increased by signal jamming.

The price process. The spectral density for price is plotted in Figure 16. Not surprisingly, the price process has significant positive serial correlation, however, there is also a lot of power at high frequencies. Thus, the price process could be viewed as the sum of a persistent process and an i.i.d. noise process. The spectral density for the full-information price process is also presented for comparison. First of all the signal-jamming spectrum is much higher than the full-information density: this reflects the fact that there is significant output, and hence variance, on the noise process in the signal-jamming case. Second, it is clear that in the signal jamming spectral density there is a significant amount of mass at high frequencies, that is, the price process is relatively noisy. But in addition it is clear that there is a major low-frequency peak, intensifying further the peak that is driven by the fundamental demand shock processes A and B. Thus, prices are far more serially correlated than are the fundamental demand shocks.

5.3. **Persistent public demand shocks and low-persistence private shocks.** I now reverse the situation considered in the previous example: I assume that the private shocks have very low persistence and that the public shocks have high innate persistence (see Table 6. Because the *private* shocks have low persistence, there is no need for firms to engage in *dynamic* signal jamming in the sense of altering the dynamic structure of the B process. See Table 7 Thus, there is very little deviation of the public output from its full-information static Cournot value of $B/3$, as is evident in the panels of Figure 18, because there is very little indirect output.

The intensity on the A process is relatively flat, so the direct output on A becomes even less serially correlated than A itself. Moreover, output on A is severely attenuated. Figure 17 shows that what little output on A there is due mostly to the indirect intensity of the rival, and so total output is driven primarily by indirect intensity from the rival.

Output on noise. Echoing these findings, it is apparent from Figure 19 that there is very little output on the noise process.

Total output. Adding these effects together in Figure 20, it is evident that total output clearly resembles the output structure of the combined full-information model, but with reduced persistence and reduced output volatility overall.

6. Conclusion

There are two main phenomena at work in this model. The first is the interaction of the two firms on the private information process. The rival firm learns about the private shocks from the history of price, and as the rival learns about the private shock, it increases its output intensity on the learned part. Concomitantly, each firm cuts back output on its own private demand shocks as the rival learns, and in the long run the rival takes over completely. ("The long run" and "at long lags" are ways to express the low-frequency elements of the model.)

The second is dynamic signal jamming on the publicly observable shock process. As the noise intensity *decreases*, information in price becomes more usable, which in turn induces more intense signal jamming output on that shock. The result is a significant reshaping of the dynamic structure of output on the public shock due to each firm's desire to disguise it. The reshaped structure can be complicated: the serial correlation of direct intensity goes down, while the serial correlation of the indirect intensity, which is the channel through which signal jamming is effected, increases. The basic numerical examples demonstrated that the serial correlation of output is raised significantly relative to the underlying serial correlation of the underlying demand shocks, and also that a significant fraction of output is driven by noise shocks, and this fraction is also highly serially correlated. The increased serial correlation is entirely due to the strategic interactions of the firms—signal jamming.

References

[1] Abreu, D., D. Pearce, and E. Stacchetti, "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory* 39, 1986, 251-269.

[2] Aghion, P., Bolton, P., Harris, C. and Jullien, B., "Optimal learning by experimentation", *Review of Economic Studies* 58, 1991, 621-654.

[3] Aghion, P., Espinosa, M., and B. Jullien, "Dynamic duopoly with learning through market experimentation," *Economic Theory*, 3(3), 1993, 517-539.

[4] Alepuz, D. and Urbano, A., "Learning in asymmetric duopoly markets: competition in information and market correlation", *Spanish Economic Review*, 7(3), 2005, 1435-5469.

[5] Amador, M. and P.-O. Weill (2012), "Learning from private and public observations of others' actions," *Journal of Economic Theory* 147 (2012) 910–940.

[6] Athey, S. and K. Bagwell, "Collusion with Persistent Cost Shocks," *Econometrica*, 2008, 76 (3), 493-540.

[7] Bagwell, K. and G. Ramey, "Oligopoly Limit Pricing", *Rand Journal of Economics*, 22, 1991, 155-172.

[8] Back, K., C. Cao, and G. Willard, "Imperfect competition among informed traders", *Journal of Finance* 55, 2000, 2117-2155.

[9] Ball, J., I. Gohberg, and L. Rodman, *Interpolation of Rational Matrix Functions* (Birkhauser Verlag, Basel, 1990).

[10] Ball, J. and B. Taub, "Factoring spectral matrices in linear-quadratic models," Economics Letters **35**, 39-44 (1991).

[11] Bergmann, D., T. Heumann and S. Morris, "Information and volatility", *Journal of Economic Theory* 158, 2015, 427-465.

[12] Bergin, J. and D. Bernhardt, "Industry Dynamics with Stochastic Demand", *Rand Journal of Economics*, 2008, 41-68.

[13] Bernhardt, D. and B. Taub, "Learning about Common and Private Values in Oligopoly," *Rand Journal of Economics*, 46(1), 2015, 66-85.

[14] Bernhardt, D., P. Seiler and B. Taub, "Speculative Dynamics," *Economic Theory*, 2010, 1-52.

[15] Bernhardt, D. and B. Taub (2006), "Kyle v. Kyle (85 vs. 89)," *Annals of Finance* 2 (1), 23-38.

[16] Bonatti, A., G. Cisterna and J. Toikka, "Dynamic oligopoly with incomplete information," *Review of Economic Studies*, 2017, 84 (2): 503-546.

[17] Caminal, R., "A Dynamic Duopoly Model with Asymmetric Information" *Journal of Industrial Economics*, 38(3), 1990, 315-333.

[18] Caminal, R. and X. Vives, "Why Market Shares Matter: An Information-Based Theory" *RAND Journal of Economics*, 27(2), 1996, 221-239.

[19] Conway, J. B., *A Course in Functional Analysis.* (Springer-Verlag, New York 1985).

[20] Davenport, W. B., Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise.* New York: McGraw-Hill (1958).

[21] Dullerud, G., and F. Paganini (2000), *A course in robust control theory.* New York: Springer.

[22] Foster, D., and S. Viswanathan, "Strategic Trading when Agents Forecast the Forecasts of Others," *Journal of Finance* LI (4), 1996, 1437-1478.

[23] Futia, Carl A. (1981), "Rational Expectations in Stationary Linear Models", *Econometrica* 49, 171–192.

[24] Green, E., and R. Porter, "Noncooperative collusion under imperfect price information," *Econometrica* 52, 1984, 87-100.

[25] Hackbarth, D., and B. Taub (2021), "Does the potential to merge reduce competition?," *Management Science*, forthcoming.

[26] Hansen, L., and T. Sargent, "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and Control* **2** (1980).

[27] Harrington, J., "Limit Pricing when the Potential Entrant is Uncertain of its Cost Function" *Econometrica*, 54(2), 1986, 429-437.

[28] Harrington, J., "Oligopolistic Entry Deterrence under Incomplete Information" *RAND Journal of Economics*, 18(2), 1987, 211-231.

[29] Harrington, J., "Experimentation and Learning in a Differentiated-Products Duopoly", *Journal of Economic Theory*, 66(1), 1995, 275-288

[30] Hoffman, K., *Banach spaces of analytic functions.* (Prentice-Hall, Englewood Cliffs 1962).
[31] Huo, Z. and N. Takayama (2021), "Rational Expectations Models with Higher Order Beliefs," mimeo, Federal Reserve Bank of Minneapolis.
[32] Kasa, K., T. Walker and C. Whiteman, "Heterogeneous Beliefs and Tests of Present Value Models", *Review of Economic Studies* (2014) 81, 1137-1163.
[33] Kasa, K. (2000): "Forecasting the Forecasts of Others in the Frequency Domain," *Review of Economic Dynamics*, 3(4), 726-756.
[34] Keller, G. and Rady, S., "Optimal Experimentation in a Changing Environment", *Review of Economic Studies*, 66, 1999, 475-507.
[35] Keller, G. and S. Rady, "Price Dispersion and Learning in a Dynamic Differentiated-Goods Duopoly," *RAND Journal of Economics* 34(1), 2003, 138-65.
[36] Klemperer, P. D., and M. A. Meyer, "Supply Function Equilibria in Oligopoly under Uncertainty" *Econometrica*, 57(6), 1989, 1243-1277.
[37] Kyle, A., "Informed Speculation with Imperfect Competition" *Review of Economic Studies* 56 1989, 317-55.
[38] Kyle, A., (1985), "Continuous Auctions and Insider Trading", *Econometrica* **53**, 1315-1335.
[39] Mailath, G. "Simultaneous Signaling in an Oligopoly Model," *Quarterly Journal of Economics*, 104(2) 1989, 417-427.
[40] Makarov, I. and O. Rytchkov (2012), "Forecasting the Forecasts of Others: Implications for Asset Pricing," *Journal of Economic Theory*, 941-966.
[41] McLennan, A., "Price dispersion and incomplete learning in the long run", *Journal of Economic Dynamics and Control* 7, 1984, 331-347.
[42] Mirman, L., L. Samuelson, and A. Urbano, "Duopoly signal jamming", *Economic Theory*, 3(1), 1993, 129-149.
[43] Nimark, K. (2017) "Dynamic higher order expectations," mimeo, Cornell.
[44] Riordan, M., "Imperfect Information and Dynamic Conjectural Variations" *RAND Journal of Economics*, 16(1) 1985, 41-50.
[45] Rondina, G., and T. Walker (2021), "Confounding Dynamics," *Journal of Economic Theory* (forthcoming).
[46] Rozanov, Y.A. *Stationary Random Processes.* Holden-Day, San Francisco (1967)
[47] Rudin, W., *Real and Complex Analysis.* (McGraw-Hill, New York 1974).
[48] Rustichini, A. and A. Wolinsky, "Learning about variable demand in the long run", *Journal of Economic Dynamics and Control* 19, 1995, 1283-1292
[49] Sannikov, Y., "Games with Imperfectly Observable Actions in Continuous Time," *Econometrica* 75 (5), 2007, 1285-1329.
[50] Sanchez-Pena, R., and M. Sznaier, *Robust Systems: Theory and Applications.* (John Wiley, New York 1998).
[51] Seiler, P. and B. Taub, "The Dynamics of Strategic Information Flows in Stock Markets," *Finance and Stochastics* 12(1), 2008, 43-82.
[52] Taub, B. (2019), "Economic and financial modelling techniques in the frequency domain," *Economic Theory Bulletin*, 7(1), 1-17.
[53] ________ (2018), "Inconspicuousness and obfuscation: How large shareholders dynamically manipulate output and information for trading purposes," *Annals of Finance* , 14(4), pp. 429-464 (2018)
[54] ________ (2009), "Implementing the Iakoubovski-Merino spectral factorization algorithm using state-space methods," *Systems and Control Letters* 58(6), 445-451.
[55] ________ "The equivalence of lending equilibria and signaling-based insurance under asymmetric information," *Rand Journal of Economics*, 21(3), 1990, 388-408.
[56] ________ (1989), "Dynamic consistency of insurance contracts under enforcement by exclusion," *Journal of Economic Dynamics and Control* 13, 93-112.
[57] ________ (1989), "Aggregate fluctuations as an information transmission mechanism," Journal of Economic Dynamics and Control 13 (1989) 113-150.

[58] Townsend, R. M. (1983), "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546-588.

[59] Vives, X., "Strategic Supply Competition with Private Information," *Econometrica*, 2011, 79, 6, 1919-1966.

[60] Whiteman, C., "Spectral Utility, Wiener-Hopf Techniques, and Rational Expectations," *Journal of Economic Dynamics and Control*, 9 (1985), 225-240.

Appendix A. Frequency-domain methods

This is an augmented version of a similar appendix that appeared in Seiler and Taub [51], which in turn built on the appendix in Whiteman [60].

Consider a serially-correlated discrete-time stochastic process a_t that can be expressed as a weighted sum of i.i.d. innovations:

$$a_t = \sum_{k=0}^{\infty} A_k e_{t-k}. \tag{53}$$

While the innovations change through time, the weights A_k remain fixed. The stochastic process can therefore be written succinctly as a function of the lag operator, L: $a_t = A(L)e_t$. The list of weights $\{A_k\}$ can be viewed as a sequence, and by the Riesz-Fischer theorem (see Rudin [47], pp. 86-90), are equivalent to functions of a complex variable z. The function of the lag operator $A(L)$ is then mathematically equivalent to a function $A(z)$ of a complex variable z. The function $A(z)$ can be analyzed with the rules of complex analysis, and this, in turn, fully characterizes the stochastic process a_t.

An important aspect of complex analysis is that the properties of a function are characterized by the domain over which they are specified. The unit disk, or sets that are topologically equivalent to the unit disk, are often the domains of interest. If a complex function on the disk can be expressed as a Taylor expansion—an infinite series where the powers of the independent variable, z, range from zero to infinity—then the function is said to be *analytic* on the disk. However, some functions, termed meromorphic functions, when expressed as a generalized Taylor expansion—a Laurent expansion—have both positive and negative powers of z, defined in an annular region containing the unit circle. This implies that they correspond to functions containing negative powers of the lag operator, which means that they operate on future values of a variable. If a variable is stochastic, this is not permissible, as it would mean that the future is predictable, contradicting its stochastic aspect. In particular, solutions to an agent's optimization problem cannot be forward-looking.

The negative powers of z in meromorphic functions arise from *poles*. The sum of the negative powers is the *principal part*.[18] To eliminate negative powers of z in a posited solution to an agent's optimization problem, we use the *annihilator operator*, $[\cdot]_+$. The annihilator operator sets the coefficients of negative powers of z in the Laurent expansion to zero, while preserving all coefficients on non-negative powers of z. This leaves a permissible, backward-looking solution to an agent's optimization problem. A function with both backward- and forward-looking parts is converted to one with only backward-looking parts by the application of the annihilator.[19]

[18]More precisely, a pole is a singularity located inside a region in the complex plane. Poles are only one possible type of singularity: there are also so-called essential singularities. Moreover, singularities need not be isolated points. In this paper the discussion focuses on rational functions, which are characterized by poles alone. Engineering terminology also refers to a function that is analytic as "causal", and the presence of poles makes it non-causal.

[19]For domain D it would be more appropriate to refer to $[\cdot]_+$ as the projection operator from $L_2(D)$ to $H^2(D)$, but the term is in widespread use.

A second property of a function concerns its invertibility.[20] If a serially-correlated stochastic process can be represented by an invertible operator, the innovations of the process can be completely and exactly recovered by observing the history of the process. That is, the inverse of the operator applied to the vector of realizations of the process yields the vector of innovations, exactly as it would if a finite vector of innovations were converted into a finite vector of realizations by an invertible matrix. A function is invertible on its domain if it does not take on a value of zero at any point inside the domain, and its inverse is then analytic. If, instead, an analytic function takes on a value of zero at a point inside the domain, then it is *noninvertible*. The inverse of a noninvertible function is not analytic. Hence, one cannot recover the vector of innovations by observing the vector of realizations, because inverting a function with a zero results in a function with negative powers of z. Recovery of the innovations would then depend on knowledge of future realizations. The *factorization* theorem of Rozanov [46] ensures that any process described by a z-transform with either negative powers of z or zeroes can be converted into an observationally-equivalent process that is characterized by an operator that is invertible and has only non-negative powers of z, so that it is backward-looking.

As an elementary example of these issues, reconsider the process in (53); if $A(L) \equiv 1-\rho L$, then the inverse operator is simply $(1-\rho L)^{-1}$, which in principle could be represented by the geometric series

$$\sum_{k=0}^{\infty} \rho^k L^k$$

but the magnitude of ρ matters for determining whether this series is convergent. When the operator $A(L)$ is translated to the frequency domain it has an equivalent representation $1-\rho z$, and in that setting convergence then becomes attached to a domain: in the example, if $|\rho| < 1$, then $1-\rho z$ is invertible on the unit disk domain because the corresponding power series for the inverse $(1-\rho z)^{-1}$, namely $\sum_{k=0}^{\infty} \rho^k z^k$, converges for any z inside the unit disk, but this does not hold outside of the unit disk. Equivalently, the pole of $(1-\rho z)^{-1}$ is ρ^{-1}, and therefore the function does not have a zero inside the unit disk and is therefore analytic there.

To illustrate the variational method, I present a simple consumer optimization problem. Consider an individual whose earnings evolve stochastically according to $y_t = A(L)e_t$, where e_t is an i.i.d., zero mean, "white noise" period innovation to earnings. The consumer's problem is to adjust bond holdings $\{b_t\}_{t=0}^{\infty}$ to maximize quadratic utility of consumption,

$$\max_{B(\cdot)} -E\left[\sum_{t=0}^{\infty} \beta^t c_t^2\right]. \tag{54}$$

subject to the budget constraint,

$$c_t = y_t + rb_{t-1} - b_t \tag{55}$$

It is possible to formulate this problem by formalising the constraint with Lagrange multipliers, but to keep the initial exposition simple, substitute the budget constraint into the

[20]In engineering parlance a function that is analytic and invertible is called minimum phase.

objective, leaving the modified problem,

$$\max_{\{b_t\}} -E\sum_{t=0}^{\infty}\beta^t(y_t + rb_{t-1} - b_t)^2, \tag{56}$$

where r is the gross interest rate satisfying $\beta r > 1$.[21] The decision problem is to choose not just the initial value of b_t, but the entire sequence $\{b_t\}_{t=0}^{\infty}$. This problem implicitly requires the choice of *functions* that react to current and possibly past states. Stationarity results in the same function applying each period.

The stochastic component of a quadratic utility function is essentially a conditional variance. If innovations are i.i.d., then the expectation of cross-products of random variables yields the sum of variances. For *white-noise* innovations, for $k > s$, $k > r$,

$$E_{t-k}\big[e_{t-r}e_{t-s}\big] = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s, \end{cases} \tag{57}$$

because of the independence of the innovations. Expressed in lag operator notation, this is

$$E_{t-k}\big[(L^r e_t)(L^s e_t)\big] = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s. \end{cases} \tag{58}$$

Notice that the "action" is in the exponents of the lag operators. From Cauchy's theorem (Conway [19]), it is equivalent to write

$$\sigma_e^2 \frac{1}{2\pi i}\oint z^r z^{-s}\frac{dz}{z} = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s, \end{cases} \tag{59}$$

where the integration is counterclockwise around the unit circle. In Cauchy's theorem, z, which is a complex number with unit radius (it is on the boundary of the disk), is represented in polar form: $z = e^{-i\theta}$. Now a more conventional integral can be undertaken, integrating over $\theta \in [0, 2\pi]$. Using Euler's theorem, which represents complex numbers in trigonometric form, $e^{-i\theta} = \cos\theta - i\sin\theta$, gives θ the interpretation of a frequency, so that z and functions of z are in the *frequency domain*.

The equivalence of (58) and (59) is crucial but might not be particularly intuitive. To see the equivalence, begin by calculating the following integral:

$$\oint_{|z|=1}\frac{1}{z}dz$$

[21] To make this problem well-defined a (small) adjustment cost must also be included, but we suppress it here because the net effect of the adjustment cost is just to make the solution stationary. Alternatively, one could simply impose the requirement that any solution be stationary.

where the integration is around the unit circle, that is, the contour integral. The the following steps demonstrate the fundamental equivalence.

$$\begin{aligned}
\oint_{|z|=1} \frac{1}{z} dz &= \oint_{|z|=1} \frac{1}{e^{i\theta}} de^{i\theta} \\
&= \int_0^{2\pi} \frac{1}{e^{i\theta}} ie^{i\theta} d\theta \\
&= \int_0^{2\pi} id\theta \\
&= i \int_0^{2\pi} d\theta \\
&= i2\pi
\end{aligned}$$

(As an aside, notice that the direction of integration around the unit-circle contour is counter-clockwise, hence it is proper to have the equivalent limits of integration in the second line as zero and 2π; clockwise integration would reverse the sign of the integral.) One can generalize this to functions of the form z^k; if $k \neq -1$,

$$\begin{aligned}
\oint_{|z|=1} z^k dz &= \oint_{|z|=1} e^{ik\theta} de^{i\theta} \\
&= i \int_0^{2\pi} e^{i(k+1)\theta} d\theta \\
&= i \left(e^{i(k+1)2\pi} - e^{i(k+1)0} \right) \\
&= i(1-1) \\
&= 0
\end{aligned}$$

Thus, defining $k = r - s$, this validates the equivalence of (58) and (59).

Whiteman [60] showed that a discounted conditional covariance involving complicated lags can be succinctly expressed as a convolution. Consider two serially-correlated processes, a_t and b_t, where

$$a_t = \sum_{k=0}^{\infty} A_k e_{t-k} \qquad \text{and} \qquad b_t = \sum_{k=0}^{\infty} B_k e_{t-k}.$$

The discounted conditional covariance as of time t, setting realized innovations to zero, is

$$E_t \left[\sum_{s=1}^{\infty} \beta^s a_{t+s} b_{t+s} \right] = E_t \left[\sum_{s=1}^{\infty} \beta^s \left(\sum_{k=0}^{\infty} A_k e_{t+s-k} \right) \left(\sum_{k=0}^{\infty} B_k e_{t+s-k} \right) \right]. \tag{60}$$

Because cross-product terms drop out, coefficients of like lags of e_t can be grouped:

$$\begin{aligned}
&\beta[A_0B_0+\beta A_1B_1+\beta^2A_2B_2+\dots]E_t[e_{t+1}^2]\\
&\qquad+\beta^2[A_0B_0+\beta A_1B_1+\beta^2A_2B_2+\dots]E_t[e_{t+2}^2]+\dots\\
&=\beta[A_0B_0+\beta A_1B_1+\beta^2A_2B_2+\dots]\sigma_e^2\\
&\qquad+\beta^2[A_0B_0+\beta A_1B_1+\beta^2A_2B_2+\dots]\sigma_e^2+\dots\\
&\qquad=\frac{\beta\sigma_e^2}{1-\beta}\sum_{s=0}^{\infty}\beta^kA_kB_k=\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint A(z)B(\beta z^{-1})\frac{dz}{z}. \qquad (61)
\end{aligned}$$

This is a useful transformation because the integrand is a product. Because the optimal policy for an optimization problem in which the objective is an expected value like that in (60), the representation in (61) permits a direct variational approach. Equation (61) is an instance of Parseval's formula, which states that the inner product of analytic functions is the sum of the products of the coefficients of their power series expansions.

A.1. **Optimization in the frequency domain.** I now apply these insights to a canonical example, the consumer's optimization problem. Hansen and Sargent ([26]) showed that the first-order conditions of linear-quadratic stochastic optimization problems could be expressed in lag-operator notation, z-transformed, and solved. Whiteman noticed that the z-transformation could be performed on the objective function itself, skipping the step of finding the time-domain version of the Euler condition.[22] The objective is then a functional, i.e., a mapping of functions into the real line. One can then use the calculus of variations to find the optimal policy function.

The first step is to conjecture that the solution to the agent's optimization problem must be an analytic function of the fundamental process e_t:

$$b_t=B(L)e_t.$$

The agent's objective (56) can then be restated in terms of the functions A and B, and the innovations:

$$\max_{B(\cdot)}-E\left[\sum_{t=0}^{\infty}\beta^t\Big((A(L)-(1-rL)B(L))e_t\Big)^2\right].$$

Expressing the objective in frequency-domain form, using the equivalence established in (61), the agent's objective can be written as

$$\max_{B(\cdot)}-\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint(A(z)-(1-rz)B(z))(A(\beta z^{-1})-(1-r\beta z^{-1})B(\beta z^{-1}))\frac{dz}{z}. \qquad (62)$$

It is immediate that a solution exists using standard methods from functional analysis.[23]

[22] A similar variational approach in continuous time can be found in Davenport and Root [20], p. 223.

[23] By reformulating the problem, the Szegö-Kolmogorov-Krein theorem (Hoffman, [30], p. 49) can be applied. The first step in this application is to re-write the argument of the integral as $|1-(1-rz)BA^{-1}|^2|A|^2$, and then re-interpret $|A|^2$ as the positive measure μ in the theorem. The second step is to transform the objective with a conformal mapping so that the transformed version of $(1-rz)$ has a zero at 0 instead of at r^{-1}; the modification of the control function $(1-rz)BA^{-1}$ then is an element of A_0, the analytic functions with a zero at 0. The Szegö-Kolmogorov-Krein theorem also provides a method for computing the value of the

A.2. **The variational method.** Let $\zeta(z)$ be an arbitrary analytic function on the domain $\{z : |z| \le \beta^{\frac{1}{2}}\}$, and let a be a real number. Let $B(z)$ be the agent's optimal choice. His objective can be restated as

$$J(a) = \max_a -\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint (A(z) - (1-rz)(B(z) + a\zeta(z)))(A(\beta z^{-1}) - (1 - r\beta z^{-1})B(\beta z^{-1}) + a\zeta(\beta z^{-1})))\frac{dz}{z}.$$

This is a conventional problem. Differentiating with respect to a and setting $a = 0$ yields the first-order condition describing the agent's optimal choice of $B(\cdot)$:

$$J'(0) = 0 = -\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint \zeta(z)(1-rz)(A(\beta z^{-1}) - (1-r\beta z^{-1})B(\beta z^{-1}))\frac{dz}{z} - \frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint \zeta(\beta z^{-1})(1-r\beta z^{-1})(A(z) - (1-rz)B(z))\frac{dz}{z}.$$

Observe the symmetry between the two integrals—everywhere βz^{-1} appears in the first integral, z appears in the second, and conversely. Whiteman establishes that the two integrals are in fact equal; we refer to this property as "β-symmetry". Therefore, the first-order condition simplifies to

$$0 = -\frac{1}{2\pi i}\oint (A(z) - (1-rz)B(z))(1-r\beta z^{-1})\zeta(\beta z^{-1})\frac{dz}{z}, \tag{63}$$

where I have dropped the leading constant $\frac{\beta\sigma_e^2}{1-\beta}$.

The integral in first-order condition (63) must be zero for arbitrary analytic functions ζ. By Cauchy's integral theorem, a contour integral around a meromorphic function with all its singularities inside the domain—a function of z that has no component that can be represented as a convergent power series expansion within the domain—is zero. Thus, all that is needed to make the integral in (63) zero is to make the integrand singular inside the disk, and to have no singularities outside the disk. The assertion is an indirect way of stating that the contour of integration is treating the outside of the circle (including ∞) as the domain over which the meromorphic function has no poles so that it is analytic there: Cauchy's theorem asserts that the integral in this sense is zero.

Recall that a solution to the agent's optimization problem must be an analytic function. The next step in the solution is to separate the forward-looking components in (63) from the backward-looking components, so that we can then eliminate the non-analytic portion from our solution. Examining equation (63), note that by construction ζ is analytic, so that it can be represented as a power series,

$$\zeta(z) = \sum_{j=0}^{\infty} \zeta_j z^j.$$

---

optimized objective, but we use a more direct approach here because we are interested in characterizing the controls themselves. I am grateful to Joe Ball for suggesting and discussing the application of this theorem with me.

This means that $\zeta(\beta z^{-1})$ has an expansion of the form

$$\zeta(\beta z^{-1}) = \sum_{j=0}^{\infty} \zeta_j \beta^j z^{-j},$$

which has only nonpositive powers of z. The negative powers of z—all but the first term—define singularities at $z = 0$, which is an element of the unit disk. However, the rest of the integrand in (63), $(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z))$, can have both positive and negative powers of z in its power series expansion. If it were possible to guarantee that only negative powers of z appeared in $(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z))$, then its expansion would take the form

$$(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z)) = \sum_{j=1}^{\infty} f_j \beta^j z^{-j},$$

for some $\{f_j\}$, and the product of this with $\zeta(\beta z^{-1})$ would take the form

$$\zeta(\beta z^{-1})(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z)) = \sum_{j=1}^{\infty} g_j \beta^j z^{-j}.$$

for some $\{g_j\}$. Every term in the sum is a singularity, and the integral of the sum is therefore zero.

The first-order condition (63) can now be broken out of the integral and stated as follows:

$$(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z)) = \sum_{-\infty}^{-1}, \tag{64}$$

where $\sum_{-\infty}^{-1}$ is shorthand for an arbitrary function that has only negative powers of z, and hence cannot be part of the solution to the agent's optimization problem. This type of equation is known as a *Wiener-Hopf equation.*

A.3. **Factorization.** To solve the Wiener-Hopf equation of a stochastic linear-quadratic optimization problem, we must factor the equation to separate the nonanalytic parts from the analytic parts. The factorization problem is a generalization of the problem of solving a quadratic equation, but there is no general formula for the solution. However, if a candidate factorization can be found, then even if it is not analytic and invertible, there is a general formula for converting that solution into an analytic and invertible factorization (Ball, Gohberg and Rodman[9]).

The Wiener-Hopf equation (64) can be restated as:

$$(1 - r\beta z^{-1})(1 - rz)B(z) = (1 - r\beta z^{-1})A(z) + \sum_{-\infty}^{-1}. \tag{65}$$

At this point it should be emphasized that the solution will be a Wiener filter, as opposed to a Kalman filter. A Kalman filter recursively reacts to information from the previous period and converges as the history of information evolves after its initiation. A Wiener filter explicitly treats history as infinite and therefore a starting date in the infinite past; the stationarity of the model dictates the use of the Wiener approach.

It is tempting to solve for $B(z)$ by dividing the left-hand side by the coefficient of $B(z)$, $(1 - r\beta z^{-1})(1 - rz)$. However, this would multiply the $\sum_{-\infty}^{-1}$ term by positive powers of z, making it impossible to establish the coefficients of the positive powers of z in the solution.

The correct procedure is first to *factor* the coefficient of $B(z)$ into the product of analytic and non-analytic functions:

$$(1 - r\beta z^{-1})(1 - rz) = \beta r^2(1 - (\beta r)^{-1}\beta z^{-1})(1 - (\beta r)^{-1}z).$$

Because by assumption $\frac{1}{r} < \beta^{1/2}$, the first factor on the right-hand side, $(1 - (\beta r)^{-1}\beta z^{-1})$, when inverted has a convergent power series (on the disk defined by $\{z \big| |z| \leq \beta^{1/2}\})$) in negative powers of z. Hence, we can divide through by this factor to rewrite the Wiener-Hopf equation as

$$\beta r^2(1 - (\beta r)^{-1}z)B(z) = \frac{(1 - r\beta z^{-1})}{1 - (\beta r)^{-1}\beta z^{-1}}A(z) + \sum_{-\infty}^{-1}, \tag{66}$$

where we use the fact that

$$\frac{1}{(1 - (\beta r)^{-1}\beta z^{-1}} \sum_{-\infty}^{-1}$$

has only negative powers of z. Because the left-hand side of (66) is the product of analytic functions, applying the annihilator to (66) yields

$$\beta r^2(1 - (\beta r)^{-1}z)B(z) = \left[\frac{(1 - r\beta z^{-1})}{(1 - (\beta r)^{-1}\beta z^{-1})}A(z)\right]_+.$$

Because $(\beta^{1/2}r)^{-1} < 1$, it follows that the inverse of $(1 - (\beta r)^{-1}z)$ is also analytic, so that we can divide by $(1 - (\beta r)^{-1}z)$ to solve for the optimal $B(z)$,

$$B(z) = \frac{\left[(1 - (\beta r)^{-1}\beta z^{-1})^{-1}(1 - r\beta z^{-1})A(z)\right]_+}{[(\beta r^2)(1 - (\beta r)^{-1}z)]}.$$

A more explicit solution for $B(z)$ obtains if the endowment process is AR(1), so that

$$A(z) = \frac{1}{1 - \rho z}.$$

Proposition 6 establishes a key result that is used repeatedly: the annihilate when there is an AR(1) construct can be simply calculated—if $A(z)$ is an AR(1), then $\left[f(\beta z^{-1})A(z)\right]_+ = f(\beta\rho)A(z)$.

Proposition 6. *("Annihilator" lemma) If f is analytic on $\beta^{-1/2}$ and $\rho < \beta^{-1/2}$, then*

$$\left[f^*(1 - \rho z)^{-1}\right]_+ = f(\beta\rho)(1 - \rho z)^{-1}.$$

Proof. Direct computation. See also Seiler and Taub [51]. □

Proposition 7 shows that the proposition about annihilates of first-order AR functions must be used with caution. If there is a zero in the annihiland, the proposition changes.

Proposition 7. *Let $a < \beta^{-1/2}$. Then* $\left[f^* \frac{1 - \frac{1}{a}z^{-1}}{1 - az}\right]_+ = 0$.

Proof.

$$\left[f^* \frac{1-\frac{1}{a}z^{-1}}{1-az}\right]_+ = \frac{1}{a}\left[z^{-1}f^*\frac{az-1}{1-az}\right]_+ = \frac{1}{a}\left[-f^*z^{-1}\right]_+ = 0.$$

□

Using Proposition 6, it follows that

$$B(z) = \frac{(1-r\beta)A(z)}{[(\beta r^2)(1-(\beta r)^{-1}\beta\rho)(1-(\beta r)^{-1}z)]}.$$

This formula has a simple "permanent income" interpretation: the agent applies the filter

$$\frac{1-r\beta}{[(\beta r^2)(1-(\beta r)^{-1}\beta\rho)(1-(\beta r)^{-1}L)]}$$

to the endowment process $A(L)e_t$ in order to smooth consumption.

A.4. **Equivalence of time domain and frequency domain approaches.** Our focus has been on generating the Wiener-Hopf equation in the frequency domain and solving it there. We now illustrate in our consumer optimization problem the general result that the time domain approach is equivalent, but less convenient.

Going back to the time domain objective in equation (56),

$$\max_{\{b_t\}} -E\sum_{t=0}^{\infty}\beta^t(y_t + rb_{t-1} - b_t)^2, \tag{67}$$

we can calculate the first order condition at time t:

$$0 = -(y_t + rb_{t-1} - b_t) + r\beta E_t[(y_{t+1} + rb_t - b_{t+1})].$$

This is an Euler equation in which the future value of the *choice variable*, b_{t+1}, appears, with the expectation of that future variable conditional on time t information. This makes the equation non-trivial in general.

The technical challenge is to calculate the expectation of the future value of the choice variable, b_{t+1}. The solution is to posit that b_t has a fixed and stationary structure, described by a filter. First recall that y_t is a serially correlated stationary process described by

$$y_t = A(L)e_t.$$

Posit that the choice variable has the stationary structure

$$b_t = B(L)e_t, \tag{68}$$

for all t. Because the conjectured structure applies to future values of the choice variable b_t, the expectation can be calculated. Note also that we made this conjecture in the development of the frequency domain approach, but *before* the optimization step.

For the conjecture to be correct, we must show that if the future and past values of b, b_{t+1} and b_{t-1}, take this form, then it is also optimal for b_t to take the same form.

Substituting the structure of y_t and the conjectured form of b_t into the Euler equation yields

$$0 = -(A(L)e_t + rB(L)e_{t-1} - B(L)e_t) + r\beta E_t[(A(L)e_{t+1} + rB(L)e_t - B(L)e_{t+1})].$$

Consolidate this further by expressing the future and lagged values of the functions using lag operators:

$$(1 - rL)B(L)e_t + r^2\beta B(L)e_t - r\beta E_t[B(L)L^{-1}e_t] = A(L)e_t - r\beta E_t[(A(L)L^{-1}e_t].$$

The next step is key. One can use the linearity of the expectation operator, and the fact that the expected value of the conditioning information is an identity, to yield

$$E_t[(1 - rL + r^2\beta - r\beta L^{-1})B(L)e_t] = E_t[(1 - r\beta L^{-1})A(L)e_t]$$

Carrying out some algebra yields

$$E_t[(1 - rL)(1 - r\beta L^{-1})B(L)e_t] = E_t[(1 - r\beta L^{-1})A(L)e_t]. \tag{69}$$

What remains is to solve this equation for B.

We have yet to specify any structure on B. However, we know that B cannot weigh future realizations of the innovations e_t: by construction they are hidden from view. However, it is possible that the general solution for B in equation (69) contains such terms. So let us posit that B has two parts, $\hat{B}$, which weights only current and past values of e_t, and $\tilde{B}$, which weights only future values of e_t, pretending for the moment that this is allowed. Substituting into (69) then yields

$$E_t[(1 - rL)(1 - r\beta L^{-1})(\hat{B}(L) + \tilde{B}(L^{-1}))e_t] = E_t[(1 - r\beta L^{-1})A(L)e_t].$$

We can isolate the $\tilde{B}$ term:

$$E_t[(1-rL)(1-r\beta L^{-1})\hat{B}(L)e_t] = E_t[(1-r\beta L^{-1})A(L)e_t] + E_t[(1-rL)(1-r\beta L^{-1})\tilde{B}(L^{-1})e_t].$$

The part on the right hand side will now be zeroed out by the expectation operator because it entails only future, unrealized and unobservable innovations. We write it suggestively as

$$E_t[(1 - rL)(1 - r\beta L^{-1})\hat{B}e_t] = E_t[(1 - r\beta L^{-1})A(L)e_t] + E_t[f(L^{-1})e_t],$$

where all we care about is that f only has terms involving L^{-1}, L^{-2}, and so on. Thus, when the expectation is taken, the result is zero; f can otherwise be arbitrary.

Removing the expectation yields

$$(1 - rL)(1 - r\beta L^{-1}) = (1 - r\beta L^{-1}) + f(L^{-1}).$$

Formally, the additional step of z-transforming the equation can now be undertaken, yielding equation (65), the same Wiener-Hopf equation obtained by taking the variational first order condition of the z-transformed objective, except that here we use the notation $f(L^{-1})$ instead of $\sum_{-\infty}^{-1}$.

As shown in the solution procedure for the frequency domain version of equation (65), this equation has a solution, validating the conjecture expressed in equation (68) that a stationary solution to the Euler equation exists. Thus, we have validated our assertion that the frequency-domain methods yield the same results as the time domain methods, that is, optimizing over optimal *quantities* in the time domain is equivalent to optimizing over *functions* in the frequency domain due to stationarity.

While the focus here has been on the familiar example of optimal consumption, all of the steps in the proof generalize: for a general problem with a quadratic objective and linear constraints, driven by stationary stochastic processes, one can complete the square of the objective, yielding a general objective of the form (62); it is helpful to express it in the equivalent form

$$\max_{\{B(\cdot)\}} -\|A - RB\|_2^2 \tag{70}$$

in which R is generically a non-invertible function. This problem is known as a model-matching problem in the engineering literature and can be solved via the generalization of the Wiener-Hopf method outlined above.

Appendix B. Derivations and proofs for the benchmark static model

Proof of Proposition 1.

Proof. Define

$$\phi(x) \equiv \left(\frac{(1+x)^2\sigma_\epsilon^2}{(1+x)^2\sigma_\epsilon^2+\sigma_a^2} \right). \tag{71}$$

Clearly, for $x \in [0,1]$, we have $T[x] \in [0,1]$. Also, if we treat ϕ as a constant, then for $x \in (0,1)$,

$$0 < |T'[x]| = \phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\frac{1}{2}\left|\frac{1}{1+x}\right|^{1/2} < \frac{1}{2}.$$

Treating ϕ as a function of x,

$$|T'[x]| = \phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\frac{1}{2}(\frac{1}{1+x})^{1/2} + \frac{\phi'(x)}{\phi}\phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\frac{1}{2}(1+x)^{1/2}$$

After algebra, we have

$$\frac{\phi'(x)}{\phi(x)} = (1-\phi)\frac{2}{1+x}.$$

Putting it all together,

$$\begin{aligned}|T'[x]| &= \phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\frac{1}{2}\left(\frac{1}{1+x}\right)^{1/2} + (1-\phi)\phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\left(\frac{1}{1+x}\right)^{1/2} \\ &= \phi^{1/2}\left(\frac{1}{2}\right)^{1/2}\left(\frac{1}{1+x}\right)^{1/2}\left(\frac{1}{2}+1-\phi\right).\end{aligned}$$

Noting that $\phi^{1/2}(\frac{3}{2}-\phi)$ achieves a maximum of $\left(\frac{1}{2}\right)^{1/2}$ when $\phi = \frac{1}{2}$, we have

$$|T'[x]| \leq \frac{1}{2}\left(\frac{1}{1+x}\right)^{1/2} \leq \frac{1}{2}.$$

which completes the argument. □

Appendix C. Dynamic model basic derivations and proofs

Proof of Proposition 2.

Proof. Conjecture that firm $-i$'s output process is a stationary linear function of its information:

$$q_{-it} = Q^{-i}(p^t+e_t;X_i^t) = \alpha_{-i}(L)A_{-i}(L)a_{-it} + \beta_{-i}(L)B(L)\overline{a}_t + \delta_{-i}(L)(p_t+e_t). \tag{72}$$

Substituting into price in (30) yields

$$\begin{aligned}\pi^P(q_{1t},q_{2t},X_t) = (1+\delta_{-i}(L))^{-1}\big[&A_1(L)a_{1t} + (1-\alpha_{-i}(L))A_{-i}(L)a_{-it} \\ &+(1-\beta_{-i}(L))B(L)\overline{a}_t - \delta_{-i}(L)e_t^P - q_{it}\big].\end{aligned} \tag{73}$$

To solve firm i's profit maximization problem, take the conjectured linear filters of firm $-i$ and the implied linear structure of prices and then optimize, solving

$$\begin{aligned}(74)\quad \max_{q_{it}} E\Bigg[\sum_{s=0}^{\infty}\eta^{t+s}\big((1+\delta_{-i}(L))^{-1}\,(A_i(L)a_{i,t+s}+(1-\alpha_{-i}(L))A_{-i}(L)a_{-i,t+s}\\ +(1-\beta_{-i}(L))B(L)\overline{a}_{t+s}-\delta_{-i}(L)e^P_{t+s}-q_{i,t+s})\big)\,q_{i,t+s}\Big|p^t+e^t,X_i^t\Bigg],\end{aligned}$$

using the structure of price from equation (73) (but leaving the price function in the conditioning information abstract to conserve notation).

Define

$$\kappa(L)\equiv(1+\delta_{-i}(L))^{-1}=\sum_{s=0}^{\infty}\kappa_s L^s, \tag{75}$$

where I assume that $(1+\delta_{-i}(L))$ is invertible,[24] and define the linear function

$$x_{it}\equiv A_i(L)a_{it}+(1-\alpha_{-i}(L))A_{-i}(L)a_{-it}+(1-\beta_{-i}(L))B(L)\overline{a}_t-\delta_{-i}(L)e_t. \tag{76}$$

Then firm i's objective can be written compactly as

$$\max_{q_{it}} E\left[\sum_{s=0}^{\infty}\eta^{t+s}\left(\kappa(L)(x_{i,t+s}-q_{i,t+s})\right)q_{i,t+s}\Big|p^t+e^t,X_i^t\right]. \tag{77}$$

The first-order condition describing firm i's best response to its rival's conjectured stationary linear strategy is

$$0=E\Big[\kappa(L)(x_{it}-q_{it})-\sum_{s=0}^{\infty}\eta^s\kappa^s q_{i,t+s}\Big|p^t+e^t,X_i^t\Big]. \tag{78}$$

The final summation captures q_{it}'s impact on future payoffs via the term $(1+\delta_{-i}(L))^{-1}$.

The linear structure of price, given the conjecture that the rival's strategy is linear, means that the price function is linear. Therefore, the conditional forecast of the net information in price in the first-order condition (78) is a linear projection on the history of (linear) prices. Thus, firm i's best response is also linear, mirroring the conjectured form for firm $-i$. Stationarity is immediate: the rival's conjectured linear strategy was not time-indexed; so the resulting linear strategy for firm i is also not time-indexed. □

[24]That is, the power series expansion of $\frac{1}{1+\delta_{-i}(L)}$ has roots inside the disk $\{z\big||z|<\eta^{-1/2}\}$ and thus is convergent. In the frequency-domain this assumption is not needed: one can manipulate objects that lack this convergence property—specifically, the zeroes of a function can lie outside the disk, but convergence is ultimately imposed by the solution procedure, specifically, the factorization step.

Proof of Proposition 2. The variational first-order conditions,[25] for α_1 and β_1 with respect to the objective (43) are:

$$\begin{aligned}
\alpha:\ \alpha_1\Big(D(1-\delta_1^* D^*)+D^*(1-\delta_1 D)\Big)A^*A\sigma_a^2 &= (D-(\delta_1+\delta_1^*)D^*D)A^*A\sigma_a^2+\sum_{-\infty}^{-1}\\
\beta:\ \beta_1\Big(D(1-\delta_1^* D^*)+D^*(1-\delta_1 D)\Big)B^*B\sigma_{\overline{a}}^2 &= (1-\beta_2)\left(D(1-\delta_1^* D^*)-D^*\delta_1 D\right)B^*B\sigma_{\overline{a}}^2+\sum_{-\infty}^{-1}
\end{aligned}$$

Rewrite these Wiener-Hopf equations as

$$\alpha:\quad \alpha_1 F^*FA^*A\sigma_a^2 = (D-(\delta_1+\delta_1^*)D^*D)A^*A\sigma_a^2+\sum_{-\infty}^{-1} \tag{79}$$

$$\beta:\quad \beta_1 F^*FB^*B\sigma_{\overline{a}}^2 = (1-\beta_2)\left(D(1-\delta_1^* D^*)-D^*\delta_1 D\right)B^*B\sigma_{\overline{a}}^2+\sum_{-\infty}^{-1} \tag{80}$$

Inverting the * terms on the left-hand side, applying the $[\cdot]_+$ operator, and inverting the remaining left-hand side coefficients yields the formulas for α and β in Proposition 2.

Derivation of the δ recursion. The next step involves developing a recursion in δ. That equation will form the initial recursion that will be developed into the recursion (48).

The δ_1 Wiener-Hopf equation is

$$\begin{aligned}
& D(1-\alpha_1)((1+\delta_2^*)D^{*2}(1-\alpha_1^*))A_1A_1^*\sigma_a^2+D(1-\alpha_2)(1+\delta_2^*)D^{*2}(1-\alpha_2^*)A_2A_2^*\sigma_a^2\\
+\ & D(1-\beta_1-\beta_2)\left((1+\delta_2^*)D^{*2}(1-\beta_1^*-\beta_2^*)\right)BB^*\sigma_{\overline{a}}^2\\
+\ & (D-1)(1+\delta_2^*)D^{*2}\sigma_e^2\\
-\ & D^{*2}(1-\alpha_1^*)(\alpha_1+\delta_1 D(1-\alpha_1))A_1A_1^*\sigma_a^2-D^{*2}(1-\alpha_2^*)\delta_1 D(1-\alpha_2)A_2A_2^*\sigma_a^2\\
-\ & D^{*2}(1-\beta_1^*-\beta_2^*)\left(\beta_1+\delta_1 D(1-\beta_1-\beta_2)\right)BB^*\sigma_{\overline{a}}^2\\
-\ & D^{*2}\delta_1 D\sigma_e^2=\sum_{-\infty}^{-1}.
\end{aligned}$$

where I have used the fact that

$$\frac{\partial}{\partial\delta_i}\delta_i D=(1+\delta_{-i})D^2.$$

This equation is fairly complicated, but significant simplification is possible because a version of the envelope theorem holds: the Wiener-Hopf equations for both α and β are embedded in the δ equation and therefore will drop out. To establish this, first divide out D^*

[25]In this setting these equations are Wiener-Hopf equations.

and bring out the factor $(1+\delta_2^*)$ to obtain:

$$\begin{aligned}
&\Big(D(1-\alpha_1)D^*(1-\alpha_1^*))A_1A_1^*\sigma_a^2 + D(1-\alpha_2)D^*(1-\alpha_2^*)A_2A_2^*\sigma_a^2\\
+\quad & D(1-\beta_1-\beta_2)D^*(1-\beta_1^*-\beta_2^*)BB^*\sigma_{\overline{a}}^2\Big)(1+\delta_2^*) + (D-1)D^*\sigma_e^2\\
-\quad & D^*(1-\alpha_1^*)(\alpha_1+\delta_1 D(1-\alpha_1))A_1A_1^*\sigma_a^2 - D^*(1-\alpha_2^*)\delta_1 D(1-\alpha_2)A_2A_2^*\sigma_a^2\\
-\quad & D^*(1-\beta_1^*-\beta_2^*)\left(\beta_1+\delta_1 D(1-\beta_1-\beta_2)\right)BB^*\sigma_{\overline{a}}^2\\
-\quad & D^*\delta_1 D\sigma_e^2 = \sum_{-\infty}^{-1}.
\end{aligned}$$

Define H by

$$\begin{aligned}
H^*H \quad \equiv \quad & (1-\alpha_1^*)(1-\alpha_1)A_1A_1^*\sigma_a^2 + (1-\alpha_2^*)(1-\alpha_2)A_2A_2^*\sigma_a^2\\
+ \quad & (1-\beta_1^*-\beta_2^*)(1-\beta_1-\beta_2)BB^*\sigma_{\overline{a}}^2 + \sigma_e^2,
\end{aligned}$$

and rewrite the Wiener-Hopf equation as

$$\begin{aligned}
D^*DH^*H\delta_1 \quad = \quad & D^*DH^*H(1+\delta_2^*) - (D^*\sigma_e^2)(1+\delta_2^*)\\
- \quad & \alpha_1 D^*(1-\alpha_1^*)A_1A_1^*\sigma_a^2 - \beta_1 D^*(1-\beta_1^*-\beta_2^*)BB^*\sigma_{\overline{a}}^2 + \sum_{-\infty}^{-1}
\end{aligned}$$

with solution

$$\begin{aligned}
(81)\quad \delta_1 = H^{-1}D^{-1}\Big[&DH(1+\delta_2^*) - H^{*-1}\Big(\sigma_e^2(1+\delta_2^*)\\
&+\alpha_1(1-\alpha_1^*)A_1A_1^*\sigma_a^2 + \beta_1(1-\beta_1^*-\beta_2^*)BB^*\sigma_{\overline{a}}^2\Big)\Big]_+
\end{aligned}$$

To isolate the α and β equations, begin by rearranging the Wiener-Hopf equation for α, equation (79) as:

$$(1-\alpha_1)\Big(D(1-\delta_1^*D^*) + D^*(1-\delta_1 D)\Big)A^*A\sigma_a^2 = D^*A^*A\sigma_a^2 + \sum_{-\infty}^{-1}.$$

Substituting for $D=(1+2\delta)^{-1}$ and $D^*=(1+2\delta^*)^{-1}$ this simplifies to:

$$(1-\alpha_1)D^*D\Big(2+\delta_1^*+\delta_1)\Big)A^*A\sigma_a^2 = D^*A^*A\sigma_a^2 + \sum_{-\infty}^{-1}. \tag{82}$$

Dividing out D^* and grouping terms yields

$$D\Big(1+\delta_2^*-\delta_1-\alpha_1(2+\delta_2+\delta_2^*))\Big)A^*A\sigma_a^2 = \sum_{-\infty}^{-1}. \tag{83}$$

Next, examine the α_1 elements in the δ_1 Wiener-Hopf equation:

$$D^*D(1-\alpha_1^*)(1-\alpha_1)(\delta_1-\delta_2^*-1) + D^*(1-\alpha_1^*)\alpha_1 A^*A\sigma_a^2.$$

Dividing out the D^* term (it appears in all of the non-α_1 terms as well) and then bringing out the common factor D yields

$$D(1-\alpha_1^*)\Big((1-\alpha_1)(\delta_1-\delta_2^*-1)+D^{-1}\alpha_1\Big)A^*A\sigma_a^2.$$

The inner terms can be rearranged to yield

$$-D(1-\alpha_1^*)\Big(1+\delta_2^*-\delta_1-\alpha_1(2+\delta_2+\delta_2^*)\Big)A^*A\sigma_a^2=\sum_{-\infty}^{-1},$$

with the last equality following from the Wiener-Hopf equation (79). Thus, these terms all drop out of the δ_1 Wiener-Hopf equation (81).

The β_1 Wiener-Hopf equation is

$$\Big((1-\beta_1-\beta_2)D^*D(2+\delta_2+\delta_2^*)-D^*(1-\beta_2)\Big)B^*B\sigma_{\bar{a}}^2=\sum_{-\infty}^{-1}. \tag{84}$$

The β terms from the δ_1 Wiener-Hopf equation are

$$\Big(D^*D(1-\beta_1^*-\beta_2^*)(1-\beta_1-\beta_2)(\delta_1-\delta_2^*-1)+D^*\beta_1(1-\beta_1^*-\beta_2^*)\Big)B^*B\sigma_{\bar{a}}^2.$$

Consolidating terms yields

$$D^*(1-\beta_1^*-\beta_2^*)\Big(D(1-\beta_1-\beta_2)(\delta_1-\delta_2^*-1)+\beta_1.\Big)B^*B\sigma_{\bar{a}}^2.$$

Adding and subtracting $1-\beta_2$ yields

$$D^*(1-\beta_1^*-\beta_2^*)\Big(D(1-\beta_1-\beta_2)(\delta_1-\delta_2^*-1)-(1-\beta_1-\beta_2)+(1-\beta_2)\Big)B^*B\sigma_{\bar{a}}^2.$$

Consolidating yields

$$D^*(1-\beta_1^*-\beta_2^*)\Big(D(1-\beta_1-\beta_2)(-2-\delta_2-\delta_2^*)+(1-\beta_2)\Big)B^*B\sigma_{\bar{a}}^2=\sum_{-\infty}^{-1},$$

with the last equality following from (84). Thus, the β elements also drop out of the δ_1 equation (81).

With the extraneous terms eliminated, the δ_1 Wiener-Hopf equation (81) reduces to

$$D^*D\Big((1-\alpha_2^*)(1-\alpha_2)A_2A_2^*\sigma_a^2+\sigma_e^2\Big)(1+\delta_2^*-\delta_1)=D^*\sigma_e^2(1+\delta_2^*)+\sum_{-\infty}^{-1}. \tag{85}$$

Next, substitute the definition of J from (45) in the main text, into equation (85) to obtain

$$D^*DJ^*J(1+\delta_2^*-\delta_1)=D^*\sigma_e^2(1+\delta_2^*)+\sum_{-\infty}^{-1}.$$

Grouping the terms above yields:

$$DJ^*J\delta_1=(DJ^*J-\sigma_e^2)(1+\delta_2^*)+\sum_{-\infty}^{-1}.$$

Solving yields

$$\delta_1 = D^{-1}J^{-1}\left[(DJ - J^{*-1}\sigma_e^2)(1+\delta_2^*)\right]_+ . \tag{86}$$

To attempt a stable recursion, further manipulation is required. We have

$$\begin{aligned}\sum_{-\infty}^{-1} &= DJ^*J(1+\delta_2^* - \delta_1) - \sigma_e^2(1+\delta_2^*)\\ &= D\Big(J^*J(1-\frac{\delta_1}{1+\delta_2^*}) - D^{-1}\sigma_e^2\Big)\\ &= D\Big(J^*J(1-\frac{\delta_1}{1+\delta_2^*}) - (1+\delta_1+\delta_2)\sigma_e^2\Big)\\ &= D\Big(J^*J(1-\frac{\delta_1}{1+\delta_2^*}) - (1+\delta_1+\delta_2)\sigma_e^2\Big).\end{aligned}$$

Now apply the annihilator operator:

$$(\delta_1+\delta_2)D = \left[-D + D\frac{1}{\sigma_e^2}J^*J(1-\frac{\delta_1}{1+\delta_2^*})\right]_+ .$$

Using symmetry and dividing by $2D$ yields

$$\delta = -\frac{1}{2} + \frac{D^{-1}}{2\sigma_e^2}\left[DJ^*J(1-\frac{\delta}{1+\delta^*})\right]_+ , \tag{87}$$

The static form of this equation exactly mirrors the static δ recursion in Bernhardt and Taub (2015).

Derivation of the recursion in D. Manipulating the definition of D in equation (87) yields

$$1+2\delta = \frac{1+2\delta}{\sigma_e^2}\left[DJ^*J\left(1-\frac{\delta}{1+\delta^*}\right)\right]_+ .$$

Dividing out $1+2\delta$ yields

$$\sigma_e^2 = \left[DJ^*J\left(1-\frac{\delta}{1+\delta^*}\right)\right]_+ . \tag{88}$$

Next undo the annihilator operator and write

$$(1+\delta^*)\sigma_e^2 + \sum_{-\infty}^{-1} = J^*J\frac{1+\delta^*-\delta}{1+2\delta}. \tag{89}$$

Now substitute

$$\delta = \frac{1}{2}(D^{-1}-1) \quad \text{and} \quad 1+\delta = \frac{1+D}{2D}$$

into the first-order condition for δ, equation (89), to obtain

$$J^*J\left(1+\frac{D^{*-1}-1-(D^{-1}-1)}{2}\right)D = \frac{1+D^*}{2D^*}\sigma_e^2 + \sum_{-\infty}^{-1}, \tag{90}$$

which reduces to

$$J^*J\left(D^*D+\frac{D}{2}-\frac{D^*}{2}\right)=\frac{1+D^*}{2}\sigma_e^2+\sum_{-\infty}^{-1}. \tag{91}$$

Starting with equation (91) we derive a new recursion in D. We first multiply equation (91) by 2:

$$J^*J\left(2D^*D+D-D^*\right)=(1+D^*)\sigma_e^2+\sum_{-\infty}^{-1}. \tag{92}$$

Next, add and subtract D:

$$J^*J\left(2D^*D+2D-(D^*+D)\right)=(1+D^*)\sigma_e^2+\sum_{-\infty}^{-1} \tag{93}$$

and then rearrange to obtain

$$2J^*JD\left(1+D^*\right)=(1+D^*)\sigma_e^2+J^*J(D^*+D)+\sum_{-\infty}^{-1}. \tag{94}$$

Divide by $2(1+D^*)$:

$$J^*JD=\frac{1}{2}\sigma_e^2+\frac{1}{2}J^*J\frac{D^*+D}{1+D^*}+\sum_{-\infty}^{-1}. \tag{95}$$

We have isolated D on the left-hand side and the D^* terms on the right-hand side, and the J terms are the compound term J^*J. After dividing by J^*, we impose the annihilator projection operator and divide by J to obtain the new recursion

$$\boxed{D=\frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+ +\frac{1}{2}J^{-1}\left[J\frac{D^*+D}{1+D^*}\right]_+}, \tag{96}$$

The next step is to express the J^*J terms in terms of D (equivalently δ).

Expressing J in terms of D. First, impose symmetry on equations (44) and (45) to obtain

$$F^*F\equiv D(1-\delta^*D^*)+D^*(1-\delta D) \tag{97}$$

$$J^*J\equiv(1-\alpha^*)(1-\alpha)AA^*\sigma_a^2+\gamma^*\gamma CC^*\sigma_c^2+\sigma_e^2. \tag{98}$$

In the existence proof in Appendix D, we will assume that $\sigma_c^2=0$, yielding

$$J^*J\equiv(1-\alpha^*)(1-\alpha)AA^*\sigma_a^2+\sigma_e^2. \tag{99}$$

To elaborate on the structure of J, it is helpful to re-express the solution for α:

$$\alpha=F^{-1}A^{-1}\left[F^{*-1}(D(1-\delta^*D^*)-D\delta D^*)A\right]_+.$$

The first step is to convert this to an expression in $1-\alpha$. Write

$$F^*F\alpha = (D(1-\delta^*D) - D^*D\delta)A + \sum_{-\infty}^{-1}.$$

Substitution from equation (97) and further manipulation yields

$$\left(D(1-\delta^*D^*) + D^*(1-\delta D)\right)A\alpha = \left(D(1-\delta^*D^*) + D^*(1-\delta D)\right)A - D^*A + \sum_{-\infty}^{-1}.$$

Bringing the common term over to the left-hand side and factoring yields:

$$(1-\alpha)\left(D(1-\delta^*D^*) + D^*(1-\delta D)\right)A = D^*A + \sum_{-\infty}^{-1}.$$

Solving yields

$$(1-\alpha) = A^{-1}F^{-1}\left[F^{*-1}D^*A\right]_+.$$

Notice that if A is a single-pole function, we can apply the annihilator lemma, Lemma 6. The annihilator term will then have the structure of A, multiplied by a constant. The leading A^{-1} term will cancel the A term inside the annihilator and thus $(1-\alpha)$ takes the form

$$(1-\alpha) = cF^{-1}, \tag{100}$$

where c is a constant. Thus,

$$(1-\alpha)AA^*(1-\alpha^*) = F^{-1}\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^*F^{*-1}, \tag{101}$$

the left-hand side of which appears in equation (99).

Again applying the annihilator lemma, assuming that A is of single-pole form, this expression becomes

$$(1-\alpha)AA^*(1-\alpha^*) = f(\eta a)^2AA^*F^{-1}F^{*-1}$$

where $f(\eta a)$ is $F(\eta a)^{-1}D(\eta a)$, reflecting the result of the annihilator lemma. To complete the derivation we need to characterize F^*F in order to characterize J^*J. We have

$$(1-\delta D) = \frac{1+\delta}{1+2\delta} = \frac{1}{2}D(1+D^{-1}) = \frac{1}{2}(1+D).$$

Therefore,

$$\boxed{F^*F = \left(D(1-\delta^*D^*) + D^*(1-\delta D)\right) = \frac{1}{2}(D^* + D + 2D^*D).} \tag{102}$$

Thus,

$$\boxed{J^*J = F^{-1}\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^*F^{*-1} + \sigma_e^2} \tag{103}$$

Equations (96), (102) and (103) and comprise a system in the functions D, F and J. These equations can then be iterated to establish many of the subsequent results.

Appendix D. Existence of equilibrium in the dynamic model

To establish existence of equilibrium in our dynamic setting, I use the recursive system in D in equation (96), showing that the associated mapping is bounded by a function that is, itself, a contraction. In our static existence argument, we assumed that cost shocks were zero, i.e., $\sigma_c^2 = 0$, and we developed a recursion in λ, proving that it was a contraction on the unit interval. If a wider domain for the recursion is allowed, the geometric approach in Bernhardt and Taub [13] (2015) shows that fixed points of the recursion can exist outside the unit interval, but the output associated with the first-order conditions evaluated at those fixed points is suboptimal for the firms. Also, the contraction property in the unit interval breaks down if the cost shock variance is too high.

As in the static model, existence fails in the dynamic model when cost shocks are too volatile, leading us to establish existence of equilibrium in the dynamic model when the cost shocks are zero. We also note that our recursion captures the restriction to the unit interval via the factorization operation: when a spectral density is factored—the generalization of taking a square root—the smaller root is automatically chosen, so that the function in question has roots inside the unit disk.

Expressing the contraction property via a variational derivative. One would like to prove that the recursion in (96) is a contraction, just as in the scalar model. In a functional recursion such as (96), a mapping μ is a contraction if there exists a positive constant $\Delta < 1$ such that

$$\frac{\|\mu(D_1) - \mu(D_2)\|}{\|D_1 - D_2\|} < \Delta.$$

When μ is a differentiable function, we can write

$$\frac{\|\mu(D_1) - \mu(D_2)\|}{\|D_1 - D_2\|} \leq \frac{\left\| \frac{\mu(D_1)-\mu(D_2)}{\|D_1-D_2\|} \right\| \|D_1 - D_2\|}{\|D_1 - D_2\|} = \left\| \frac{\mu(D_1) - \mu(D_2)}{\|D_1 - D_2\|} \right\| \sim \left\| \frac{\partial}{\partial D}\mu(D) \right\|,$$

where the derivative is the variational derivative. The result is the norm of the derivative, not the derivative of the norm. To develop intuition, we verify that this condition holds in a simplified quasi-scalar version of the model, using a conventional derivative rather than a variational derivative.

Intuition from scalar case. The dynamic model reduces to the scalar model when the persistence parameters b, ρ and ϕ are zero. Intuition about the contraction property can be gleaned by considering the ordinary derivative of a scalar version of (96). Then, D, F and J become ordinary real variables, not functions of z, so the annihilator operator becomes the identity, $D^* = D$, etc.

We first analyze the second term in the D recursion (96): in the scalar version of $J^{-1}\left[J\frac{D^*}{1+D^*}\right]_+$, the annihilator operator is not present, leaving $\frac{1}{2}\frac{2D}{1+D} = \frac{D}{1+D}$. The derivative is

$$\frac{d}{dD}\frac{D}{1+D} = \frac{1}{(1+D)^2} < 1,$$

as long as D is strictly positive.

Now consider the first term in (96). In a scalar setting, substituting from (102), equation (101) becomes

$$
\begin{aligned}
(1-\alpha)AA^*(1-\alpha^*) &= F^{-1}\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^* F^{*-1}\\
&\sim (D^*D)(F*F)^{-2}\frac{1}{(1-a)^2}\\
&\sim \frac{D_n^2}{(\frac{1}{2}(2D+2D^2))^2}\frac{1}{(1-a)^2}\\
&= \frac{D^2}{\frac{1}{4}(2D+2D^2)^2}\frac{1}{(1-a)^2}\\
&= \frac{1}{(1+D)^2}\frac{1}{(1-a)^2}
\end{aligned}
\tag{104}
$$

where we arbitrarily write the scalar value of A as $\frac{1}{1-a}$. The recursion equations (103) and (96) then become a difference equation system,

$$J_{n+1}^2 = \frac{1}{(1+D_n)^2}\frac{1}{(1-a)^2}\sigma_a^2 + \sigma_e^2 \tag{105}$$

and

$$D_{n+1} = \frac{1}{2}\frac{1}{J_n^2}\sigma_e^2 + \frac{D_n}{1+D_n}. \tag{106}$$

Notice from the definition of J in equation (45) that the first term is bounded, i.e.,

$$\frac{1}{2}\frac{1}{J_n^2}\sigma_e^2 \le \frac{1}{2},$$

and $J(0) \neq 0$, implying that it is not a fixed point.

We can analyze the nonlinear system (105) and (106) for stability. For large values of D_n, J_{n+1} is approximately σ_e, so that D_{n+1} is driven to approximately $\frac{1}{2}+1$. For very small values of D_n, J_n approaches a constant, and therefore D_{n+1} also approaches a constant. Moreover, this fixed point is stable, as (setting σ_a^2 to one and a to zero for simplicity) the derivative

$$\frac{d}{dD_n}\frac{1}{2}\frac{1}{J_n^2}\sigma_e^2 = \sigma_e^2\frac{\frac{1}{(1+D_n)^3}}{\left(\frac{1}{(1+D_n)^2}+\sigma_e^2\right)^2} = \frac{\frac{1}{(1+D_n)^2}}{\left(\frac{1}{(1+D_n)^2}+\sigma_e^2\right)}\frac{\sigma_e^2}{\left(\frac{1}{(1+D_n)^2}+\sigma_e^2\right)}\frac{1}{(1+D_n)}$$

is obviously a fraction if D_n is positive. Thus, if the initial value of D_n is positive, this (scalar) recursion is stable and has a positive fractional fixed point.

Existence proof for the dynamic model. The main recursion, equation (96), is complicated by the presence of the annihilator operator. Were the annihilator operator not there, we could execute a direct proof of the contraction property. However, the annihilator operator necessitates an indirect approach. The indirect approach entails finding an ancillary mapping T that (1) bounds the mapping S implicitly defined by the right hand side of (96), and (2) is itself bounded and a contraction. The ancillary mapping is tractable, so it is straightforward to characterize the domain over which it is a contraction. We show that S also maps this domain into itself and is continuous. It then follows that a fixed point of S exists.

Lemma 4. *Let X be a Banach space. Let $T : X \to X$ and $S : X \to X$ be mappings such that*

(i) *T is bounded and a contraction;*
(ii) *S is continuous with $\|S\| \leq \|T\|$ on a compact and convex subset $\overline{X}$ of X that includes the fixed point of T.*

Then a fixed point of S exists in $\overline{X}$.

The space X in our setting is a Hardy space $H^2[\eta]$, that is, the space of square integrable functions on the η disk, i.e., the elements z in the complex plane such that $\{z \,\big|\, |z| \leq \eta^{-1/2}\}$. The function D, which is our object of interest, is an element of X. The space $H^2[\eta]$ is a Hilbert space, and as such is a complete normed vector space, and as such is a Banach space.[26] Because it is a Banach space we can establish that there is a fixed point by invoking Schauder's fixed point theorem.

Lemma 4 does not deliver uniqueness of the fixed point. However, we conjecture that the fixed point and associated equilibrium are, in fact, unique (given sufficiently little uncertainty about private values).

Proof. The sole issue is to identify the compact subset $\overline{X}$. We define the set using the contraction property. Let x^* be the fixed point of T. Let

$$X_0 \equiv \{x : 0 \leq |x| \leq |x^*|\}\,.$$

This set is closed and bounded. The upper bound of $|T[X_0]|$ is finite due to the contraction property. The upper bound of $T[T[X_0]]$ is also finite, and by the contraction property must be closer to the fixed point $|x^*|$; and this holds for all iterations $T[\ldots T[T[x]]\ldots]$. Define

$$\overline{X} \equiv \{x : 0 \leq |x| \leq \sup|T[X_0]|\}\,,$$

which is a closed and bounded (compact) set and trivially convex. Because $\|S\| < \|T\|$, and because $T[\overline{X}] \subseteq \overline{X}$ by the contraction property, $S[\overline{X}] \subseteq \overline{X}$. Because S is a continuous mapping, we can apply Schauder's fixed point theorem to establish that a fixed point of S exists. □

To apply Lemma 4, we first show that our recursion satisfies its key inequality, i.e., there is a bounding mapping T that is a contraction. Viewing D as an element of $H^2[\eta]$, define the mapping:

$$T[D] \equiv \frac{1}{2} + \frac{D}{1+D}. \tag{107}$$

Also define the mapping associated in the recursion in (96) by S. We begin with:

[26]See Seiler and Taub (2008), Appendix C for properties of $H^2[\eta]$.

Lemma 5.

$$|D| = |S[D]| = \left|\frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+ + \frac{1}{2}J^{-1}\left[J\frac{2Re[D^*]}{1+D^*}\right]_+\right| \le \frac{1}{2} + \left|\frac{D}{1+D}\right|. \tag{108}$$

That is, $|S| \le |T|$.

Proof. The first term is easy. The absolute value (and therefore the norm) passes through the annihilator operator (see the appendix of Seiler and Taub 2008):

$$\left|\frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+\right| \le \frac{1}{2}\left|J^{-1}\right|^2\sigma_e^2 = \frac{1}{2}\left|J^{-1}J^{*-1}\right|\sigma_e^2 \le 1$$

by the construction of J. For the second term, we have

$$\begin{aligned}\frac{1}{2}\left|J^{-1}\left[J\frac{D+D^*}{1+D^*}\right]_+\right| &\le \frac{1}{2}\left|J^{-1}\right|\left|J\right|\left|\frac{D+D^*}{1+D^*}\right| \\ &\le \frac{1}{2}\left|\frac{D+D^*}{1+D^*}\right| \\ &\le \frac{1}{2}\left|\frac{D^*\left(1+\frac{D}{D^*}\right)}{1+D^*}\right| \le \frac{1}{2}\left|\frac{D^*}{1+D^*}\right|\left|\left(1+\frac{D}{D^*}\right)\right| \\ &\le \left|\frac{D^*}{1+D^*}\right| \le \left|\frac{D}{1+D}\right|.\end{aligned}$$

□

Note that the cancellation of J and J^{-1} would not necessarily work were we calculating the sup norm instead of the absolute value at the same value of z.

The next lemma establishes that the bound mapping T is contractive.

Lemma 6. *If the domain of T is such that $|1+D|>1$, then T is a contraction and T is bounded.*

Proof. The final term of T becomes contractive: the variational derivative is

$$\left|\frac{\partial}{\partial D}\frac{D^*}{1+D^*}\right| = \left|\frac{1}{(1+D^*)^2}\right| < 1, \quad \text{when } |1+D|>1.$$

□

To establish boundedness, we must prove that $|1+D|>1$ in the vicinity of the fixed point. We do this in Proposition 3.

Lemma 7. *S is a continuous mapping.*

Proof. Because the recursion is nonlinear, we establish continuity component by component: the elements of S include inversion (J^{-1}), the annihilator operator ($[\cdot]_+$), factorization (J), and the construction of J, which involves D nonlinearly. We must show that each of these elements preserves continuity. To show that J is a continuous function of D, we use the Szegö factorization. The Szegö factorization is the generalization of representing a function in exponential-log form: for a function $f(x)$, we can write $e^{\ln f(x)}$. If the function f is a function of a complex variable and is two sided, i.e., $f(z) = A(z)A(z^{-1})$, then the Szegö form allows one to effectively take the square root and recover $A(z)$. One can then indirectly demonstrate properties of the function $A(z)$.[27]

Using the Szegö form for the J function, we can write

$$J(\alpha) = e^{\frac{1}{2}\frac{1}{2\pi i}\oint \frac{\zeta+\alpha}{\zeta-\alpha}\ln(J^*(\zeta)J(\zeta))\frac{d\zeta}{\zeta}}.$$

Because the exponential function is continuous, we just need to show that J^*J is continuous in D.

The annihilator operator can be expressed with the Szegö form,

$$\left[J^{*-1}\right]_+ = J(0)^{-1} = e^{\frac{1}{2}\frac{1}{2\pi i}\oint \ln\left(J^{*-1}J^{-1}\right)\frac{d\zeta}{\zeta}},$$

as can the inverse $J(z)^{-1}$,

$$J^{-1}(z) = e^{\frac{1}{2}\frac{1}{2\pi i}\oint \frac{\zeta+z}{\zeta-z}\ln\left(J^*(\zeta)^{-1}J(\zeta)^{-1}\right)\frac{d\zeta}{\zeta}}.$$

However, because $|z| = 1$ and in the Szegö factorization, $|\alpha| < 1$, this expression holds in the limit.

Recalling equations (102) and (103),

$$J^*J = 2\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^*(D^*+D+2D^*D)^{-1}\sigma_a^2 + \sigma_e^2 \tag{109}$$

and we just need to establish continuity for this object. $(D^*+D+2D^*D)^{-1}$ is continuous in D for $D > 0$. We can also establish that $\left[F^{*-1}D^*A\right]_+$ is continuous in D by using the Szegö factorization, but because of the annihilator lemma we calculate it at a (recall that $A = \frac{1}{1-az}$):

$$\left[F^{*-1}D^*A\right]_+ = F(a)^{-1}D(a)A(z) = e^{-\frac{1}{2}\frac{1}{2\pi i}\oint \frac{\zeta+a}{\zeta-a}\ln\left(D^*D(D^*+D+2D^*D)^{-1}\right)\frac{d\zeta}{\zeta}}A(z) \tag{110}$$

which is continuous due to the continuity of the product, exponential, and $(D^* + D + 2D^*D)^{-1}$. □

To apply Lemma 4 we show that $D = 0$ is not a fixed point of the recursion S.

Lemma 8. $J(0) \neq \infty$.

Proof. Use the Szegö factorization to write

$$F^{-1}\left[F^{*-1}D^*A\right]_+ = e^{-\frac{1}{2}\frac{1}{2\pi i}\oint \frac{\zeta+a}{\zeta-a}\ln\left(D^*D(D^*+D+2D^*D)^{-2}\right)\frac{d\zeta}{\zeta}}A(z). \tag{111}$$

The inner term can be written as

$$D^*D(D^*+D+2D^*D)^{-2} = \frac{D}{(\frac{D}{D^*}+1+2D)}\frac{1}{(D^*+D+2D^*D)} = \frac{1}{(\frac{D}{D^*}+1+2D)}\frac{1}{(\frac{D^*}{D}+1+2D^*)}.$$

$\frac{D}{D^*}$ and $\frac{D^*}{D}$ are bounded away from zero (to see this, express D in polar form). Therefore, the whole denominator is bounded away from zero at $D = 0$. Thus, $J(0)$ is finite. □

[27]See Taub [55] (1990) for a more thorough discussion of the Szegö form.

Lemma 9. *$D = 0$ is not a fixed point of the bounding function T.*

Proof. Substitution yields $T[0] = \frac{1}{2}$. □

To complete the argument, we find a positive *lower* bound for the mapping S, i.e., a bound $\underline{D}$ such that if $|D_1| > \underline{D}$ then $|S[D_1]| > \underline{D}$, so that any fixed point is then bounded away from zero.

Lemma 10. *There exists a lower bound $\underline{D}$ such that $|S[D]| > \underline{D}$ for all D.*

Proof. From Lemma 8, we have $\left|\frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+\right| \geq \frac{1}{2}\frac{\sigma_e^2}{\sigma_e^2} = \frac{1}{2}$. □

Having determined a lower bound we can combine this with the upper bound induced by the mapping T, leading to the following corollary:

Corollary 2. *There is a $\xi > 0$ with*

$$X_\xi \equiv \{D : \xi|D| < 1\}$$

such that for $D \in X$,

$$S[D] \in X_\xi.$$

We now have the ingredients to assert

Proposition 8. *A fixed point of S exists.*

Proof. The contraction property for T requires $|1 + D| > 1$:

$$(1 + D)(1 + \overline{D}) = 1 + 2\mathrm{Re}(D) + |D|^2 > 1,$$

so

$$\mathrm{Re}(D) > -\frac{|D|^2}{2},$$

which is satisfied by $\mathrm{Re}(D) > 0$. If we define X_0 as the smaller set

$$X_0 \equiv \{D : \mathrm{Re}(D) > 0\}$$

we satisfy this requirement. Then we just need to show that if $\mathrm{Re}(D) > 0$, then $\mathrm{Re}(T[D]) > 0$, i.e.,

$$\frac{1}{2} + \mathrm{Re}\frac{D}{1 + D} > 0.$$

This follows because the denominator of $\frac{D}{1+D}$ has a larger real part than the numerator, but the same imaginary part. This means that if we represent D in polar form, $D = D_0 e^{i\theta}$, then $1 + D$ will have the form $\tilde{D}_0 e^{i\tilde{\theta}}$, where $\left|\tilde{\theta}\right| < |\theta|$ and $\tilde{D}_0 > D_0$. In expressing this in geometric form it is evident that $\mathrm{Re}\frac{D}{1+D} > 0$ and therefore that $\frac{1}{2} + \mathrm{Re}\frac{D}{1+D} > 0$. Thus, T maps X_0 into X_0, and in addition $|1 + D| > 1$ and the contraction property holds for T as defined in equation (107). The properties listed in Lemma 4 are satisfied for S and T by Lemmas 5, 6, 7.

It remains to verify that the fixed point reflects an optimum, i.e., that the associated solutions of the Wiener-Hopf equations for α, β, and δ are optimal. Consider α. Inspection of the objective (43) and the variational first-order condition (79), reveals that the variational second-order condition for α is

$$-\oint F^* F A^* A \sigma_a^2 \frac{dz}{z} = -\|FA\|_2^2 < 0.$$

Thus, the solution for α in equation (46) represents an optimum. The optimality of β and δ follow similarly. □

Proposition 3 in the main text follows.

Appendix E. Proofs of the characterisation results

Before proving Proposition 4 I begin with a series of preliminary lemmas. The first lemma establishes a basic property of signal extraction. I then prove that this property is violated in equilibrium. Finally, I prove that outputs are not just scalar amplifications of the input shocks.

Consider a first-order autoregressive (AR) process

$$x_t = A(L)e_t = \frac{1}{1-\rho L} e_t.$$

Because I will treat the problem in terms of poles, write this as

$$-\frac{\rho^{-1}}{L-\rho^{-1}} e_t,$$

where ρ^{-1} is the pole. I suppose that this process cannot be observed directly, but that there is an observable signal process

$$y_t = A(L)e_t + u_t,$$

where u_t is a white noise process, uncorrelated with e_t.

The signal extraction problem is to construct a filter $F(\cdot)$ that optimally extracts information from this noisy signal, producing an output process $F(L)(A(L)e_t + u_t)$:[28]

Lemma 11. *The poles of the signal extraction output process are the same as the poles of the input process. Signal extraction is expressed entirely in the moving average part of the filtered process.*

Proof. I use frequency-domain methods. I solve the optimal filtering problem

$$\begin{aligned}(112)\quad & \min_F E(A(L)e_t - F(L)(A(L)e_t + u_t))^2 \\ & = \min_F \frac{1}{2\pi i} \oint \Big((A-FA)^*(A-FA)\sigma_e^2 + F^* F \sigma_u^2 \Big) \frac{dz}{z}.\end{aligned}$$

The variational first-order condition is

$$A^*(A-FA)\sigma_e^2 - F\sigma_u^2 = 0.$$

[28]This result was stimulated by a personal exchange with Ken Kasa and Charles Whiteman.

The right hand side is zero instead of $\sum_{-\infty}^{-1}$ because the filter is allowed to be two-sided. The solution is

$$F = (A^* A \sigma_e^2 + \sigma_u^2)^{-1} A^* A$$

The poles of A completely cancel, leaving an ARMA part where the poles (the denominator part) come from the MA part of the noisy process. When one hits the noisy process with this filter, the MA part of the noisy process cancels, but the forward-looking part of the filter's poles remains.

Repeating the process with a one-sided filter, the variational first-order condition is

$$A^*(A - FA)\sigma_e^2 - F\sigma_u^2 = \sum_{-\infty}^{-1}$$

or

$$-(A^* A \sigma_e^2 + \sigma_u^2)F + A^* A \sigma_e^2 = \sum_{-\infty}^{-1}.$$

Define the factor H by

$$H^* H = A^* A \sigma_e^2 + \sigma_u^2.$$

The poles of H are the same as the poles of A. The solution is

$$F = H^{-1} \left[H^{*-1} A^* A \right]_+ .$$

Recall that if some function $A(z)$ is an AR(1), then by Lemma 6, $[f^* A(z)]_+ = f(\eta\rho)A(z)$. Using the assumption that A is the sum of AR(1) terms—that is, that the number of poles of A exceeds the number of zeroes, and using the linearity of the annihilator operator, one can apply this fact term by term, with the result that the poles of the annihilate $\left[H^{*-1}A^*A\right]_+$ are the same as the poles of A. The poles of H, which are the zeroes of H^{-1}, then cancel the poles of the annihilate. □

When one hits the noisy process—which is characterized by H—with this filter, the H parts cancel, leaving a sum of AR's, but weighted differently from the original A process. The numerator of the filtered process—the MA part—has the signal extraction information. Importantly, there are no new poles; the original poles, and only those poles, are preserved in the product FH.

Thus, were signal extraction the only force determining the output process, the poles of the input process would be preserved in the output process, and there would be no new poles. We are now prepared to prove Proposition 4 that equilibrium output intensity filters α_i and β_i are not scalar-valued—output intensities are not just amplifications of the dynamic shock processes.

Proof. **(of Proposition 4)** Because the exogenous shock processes are first-order autoregressive (AR(1)) processes, the frequency domain filters A and B have single poles. Suppose by way of contradiction that the intensities α_i and β_i, and δ_i are all scalar. Then, write equation (46) (the equation for α_i), as

$$\begin{aligned}\alpha_i &= F^{-1}A^{-1}\left[F^{*-1}(F^*F - D^*)A\right]_+ \\ &= F^{-1}A^{-1}\left([FA]_+ - \left[F^{*-1}D^*A\right]_+\right).\end{aligned} \tag{113}$$

The annihilator operator $[\cdot]_+$ is an identity in the first term on the right hand side. In the second term because the $A(\cdot)$ function is a single pole form, the projection operator $[\cdot]_+$ yields a constant multiplying $A(\cdot)$ (from the "annihilator lemma", Lemma 6) . The $A(\cdot)$ function is then canceled by the $A^{-1}(\cdot)$ term, leaving the right hand side as a pure scalar if F^{-1} is scalar. Thus, α_i is a scalar if F^{-1} is a scalar. Similar reasoning applies in the β_i equation.

For F^{-1} to be a scalar, F must be scalar. The definition of F in equation (44) reveals that F is scalar only if δ and D are scalar, which is true by our maintained assumption. For D to be scalar, J would need to be scalar. But J cannot be scalar: in equation (47), a scalar F and D means that J is driven by the A filter, which is exogenously non-scalar, a contradiction. □

Appendix F. State space methods in the numerical analysis

In order to numerically simulate and iterate the recursion in equation (48), I constructed algorithms using so-called state space methods from the control systems engineering literature. These methods suppose that the stochastic processes in a system have an ARMA (autoregressive-moving average) structure, but can otherwise be arbitrary vector processes, that is, a process can be represented as

$$x_t = Ax_{t-1} + Bu_t \tag{114}$$

where x_t and u_t can be vector processes, and A and B are appropriately conformable matrices. In engineering settings the x_t process would be considered the state process, and the u_t process would be a serially uncorrelated and i.i.d. process, that is, white noise. When x_t and u_t are scalar-valued and A and B are scalar constants, this is simply an AR(1) process. Intuitively, for an AR(1) process to be stable requires that $|A| < 1$; this stability notion generalizes: a more general vector-valued system is stable if the eigenvalues of A are less than one in absolute value.

There might be an output process driven by this state,

$$y_t = Cx_t + Du_t \tag{115}$$

where y_t can also be a vector process, and C and D are again appropriately conformable matrices. For example, y_t might be the observation of a noisy state that one would want to estimate using Kalman filter methods.

We can write (114) using the lag operator L:

$$x_t = ALx_t + u_t$$

and if A has the appropriate structure, namely eigenvalues less than one, we can solve:

$$x_t = (I - AL)^{-1} u_t.$$

Substituting into (115) yields

$$y_t = (C(I - AL)^{-1} B + D) u_t$$

that is, the output process is expressed entirely in terms of the underlying fundamental or driving process u_t. This is simply the generalization of an ARMA, not just an AR, process.

It is now convenient to use that fact that was developed in Appendix A, namely that the lag operator maps into a element of the complex plane, which we denote z, and we can represent the process y_t simply by its z-transform,

$$C(I - Az)^{-1} B + D \tag{116}$$

This expression is the generalization of a rational function.

It is convenient to re-express models of this type with the *inverse* of the A matrix, that is,

$$C(zI - A)^{-1} B + D \tag{117}$$

and the eigenvalues of A now need to *exceed* one for stability to hold. This engineering convention will be used in the exposition from this point forward; the eigenvalues are then referred to as the *poles*. A process expressed in this way is a state space *realization*.

Importantly, the realization form in expression (117) is preserved when familiar algebraic operations are carried out on the expression. For example, the sum of two processes that are constructed from the same driving process u_t can be expressed as

$$(C_1(zI-A_1)^{-1}B_1+D_1)+(C_2(zI-A_2)^{-1}B_2+D_2) = \begin{pmatrix} C_1 & C_2 \end{pmatrix} \left(zI - \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \right)^{-1} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} + \begin{pmatrix} D_1 + D_2 \end{pmatrix} \tag{118}$$

which has the same basic form as (117). Because the form is preserved, the engineering literature has developed a special notation for it:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

The addition operation can be expressed in this notation by

$$\left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array} \right].$$

Similarly, multiplication and inversion are expressed as

$$\left[\begin{array}{cc|c} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1 C_2 & D_1 D_2 \end{array} \right].$$

$$\left[\begin{array}{c|c} A - BD^{-1}C & BD^{-1} \\ \hline -D^{-1}C & D^{-1} \end{array} \right]$$

The details of these and other operation can be found in Dullerud and Paganini [21], p. 99, or in Sanchez-Pena and Sznaier [50], p. 465-470. Other operations such as transposition and complex conjugation are also straightforward.

There are two other operations that can be expressed using state space methods: annihilation, that is, the annihilation operator $[\cdot]_+$ that was discussed in Appendix A, and spectral factorization. All of these operations—addition, multiplication, conjugation and transposition, annihilation, and spectral factorization—are used in the recursion equation (48).

Finally, it is possible to numerically calculate norms using the realization by solving a Lyapunov equation; see p. 475 of Sanchez-Pena and Sznaier [50] (1998)).

The realization for a system is not necessarily unique. Specifically, we can construct transformations of a realization to manipulate the A matrix, that is, we can calculate

$$CT^{-1}(zI - TAT^{-1})^{-1}TB + D = \tilde{C}(zI - \tilde{A})^{-1}\tilde{B} + D$$

Such transformations can usefully isolate characteristics of the system, and importantly, provide ways of it approximating the system with a smaller system, that is, one in which the dimension of the A matrix is reduced: this is called balanced truncation. In balanced truncation, the so-called controllability and observability Gramians are calculated via solving a Lyapunov equation for each. A coordinate transformation is chosen so that these Gramians are identical. The singular values of the Gramians—the square roots of the eigenvalues—then can be ordered, with the largest corresponding to the sup norm of the system in question. The elements of the system associated with the smallest singular values can then be discarded, if the resulting change in the infinity norm of the resulting system is dominated by the chosen tolerance. This reduces the number of poles. Moreover, the error entailed in the reduction of the model has an analytical bound that is a linear function of the sums of the discarded singular values. We use the balanced truncation algorithm of Laub and Glover (see p. 319 of Sanchez-Pena and Sznaier (1998)).

There is an additional operation needed in the numerical calculations: minimal realization. A minimal realization generalizes the idea of canceling the poles and zeroes of a rational function if they are equal. Thus, if we are given a rational function

$$\frac{(1 - .2z)(1 - .3z)}{(1 - .2z)(1 - .7z)}$$

it is obviously equivalent to

$$\frac{(1 - .3z)}{(1 - .7z)}$$

but what about

$$\frac{(1 - .2z)(1 - .3z)}{(1 - .2001z)(1 - .7z)}?$$

The state space approach generalizes rational functions of this sort. As a result of the other approximations that are carried out from operations such as inversion, spectral factorization, and balanced truncation, small numerical errors can make the coefficients in the numerator and denominator that should cancel slightly different; minimal realization algorithms force the cancellation if the coefficients satisfy a tolerance.

We use the Kung algorithm to compute the minimal realization (see p. 310 of Sanchez-Pena and Sznaier (1998)). The algorithm sets up a block Hankel matrix of the system and uses singular value decomposition (a generalization of diagonalization of a matrix) to factor the Hankel matrix. The tolerance level removes nearly zero singular values, so that the remaining system is both controllable and observable—which translates into pole-zero cancellation when there is numerical noise.

The algorithm. I implemented these operations using Mathematica in order to numerically approximate a fixed point of equation (48), the algorithm works as follows:

(1) An initial conjecture of the solution of $D(z)$ is posited (not to be confused with the notation D for the state-space realization);
(2) This conjecture is used in equation (102) where a spectral factorization is carried out to calculate F, using the method devised in Taub [54], and in turn the calculated value of F as well as the conjectured value of $D(z)$ is used in the spectral factorization in equation (103) to calculate J;
(3) The resulting value of J and the conjectured value of D are substituted on the right hand side of the recursion (48) and the requisite multiplication, inversion, annihilation and addition operations are carried out, resulting in a new value of D, which becomes the new conjecture;
(4) The iteration terminates when a Cauchy-style convergence criterion is met, that is, for iteration i, the norm of the improvement $\|D_i - D_{i-1}\|_2$ falls below the chosen tolerance.

There are some further details of the algorithm that bear mention. Examining equation (48), it is apparent that there are some inverses in the equation, as well as some spectral factorizations. These inversions and factorizations increase the number of pole terms on each iteration. The balanced truncation algorithm trims the insignificant pole terms in the iteration.

The spectral factorization algorithm must cope with the arbitrary number of pole terms that arise from the proliferation of poles in the iteration. For this reason a more robust spectral factorization algorithm is needed, and this is provided by the algorithm in Taub [54]. This procedure also requires the choice of a tolerance.

There are thus four tolerances that must be chosen to run the algorithm: the spectral factorization tolerance, the balanced truncation tolerance, the minimal realization tolerance, and the Cauchy criterion for D. Excessively relaxing the tolerances leads to unstable behavior numerically. When appropriate tolerances are chosen, the system converges numerically, as is predicted by the contraction property established in Appendix D.

Appendix G. A primer on the algebra of spectral densities

Much of our analysis rests on the use of spectral densities. Spectral densities have some algebraic properties that we review here. Before beginning, we note that the spectral density is a function, but the essential characteristics of the density can be obtained by plotting the density.

The spectral density is defined simply as the absolute value of a function of z, but using the definition $z = e^{-i\theta}$:

$$|f|^2 = f(z)f(z^{-1}) = f(e^{-i\theta})f(e^{i\theta}) \tag{119}$$

where we plot this as a function of θ. The spectral density can be thought of as the localized variance of a stochastic process characterized by the filter $f(z)$. The overall variance is then just the integral under the spectral density. It is identical to the H_2 norm of the process, obtained via the contour integral of the absolute value squared of the filter.

The first key aspect of the spectral density is that it reflects the serial correlation of the underlying stochastic process. Figure 1 plots the spectral density for a positively correlated moving average process $1 + .5z$ and a positively serially correlated first-order autoregressive process $(1 - .5z)^{-1}$. Both plots are positively bowed, but the autoregressive process has a peak in the middle, reflecting the long run persistence of the process. The middle part of the plot corresponds to low frequencies (θ near zero) and the outer shoulders correspond to the high frequencies.

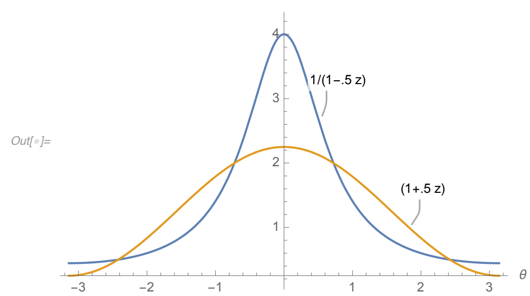

FIGURE 1. Spectral densities for positively serially correlated processes.

Increased persistence in the underlying process increases the sharpness of the central peak (see Figure 2):

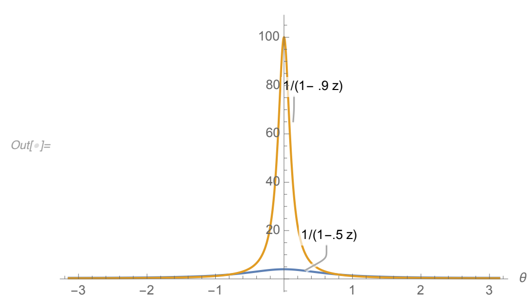

FIGURE 2. Spectral densities for positively serially correlated processes.

Negative serial correlation has the opposite pattern. Figure 3 plots the spectral density for a negatively correlated moving average process $1 - .5z$ and a negatively serially correlated first-order autoregressive process $(1 + .5z)^{-1}$. Both plots are negatively bowed, but the autoregressive process has a peak in the middle, reflecting the long run persistence of the process. The middle part of the plot corresponds to low frequencies (θ near zero) and the outer shoulders correspond to the high frequencies.

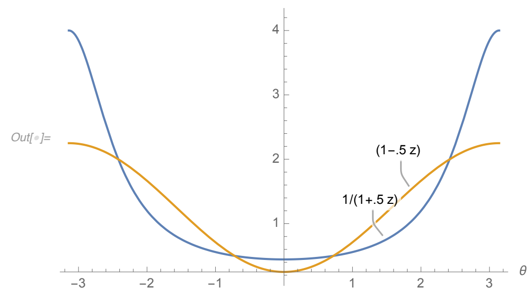

FIGURE 3. Spectral densities for negatively serially correlated processes.

Now the long run serial correlation of the autoregressive process is reflected by the higher power in the right and left shoulders of the plot.

The next key characteristic of spectral densities concerns the arithmetic operation of multiplication. For two filters f and g then the spectral density of the product fg filter will be the the product of the two separate densities at each frequency; see Figure 4. Thus, for

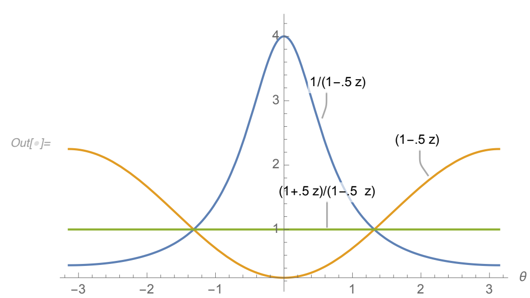

FIGURE 4. Spectral density of a product.

example, firms determine output by applying an intensity filter to the fundamental serially correlated demand shock process; the spectral density of the resulting output process will in general be very different from the respective densities of the demand process and the intensity filter.

The next point concerns addition. The spectral density of a sum is *not* the sum of the spectral densities for two filters operating on the same underlying fundamental process. Figure 5 displays the spectral density for the sum of two of the filters we have seen before. Figure 6 displays the spectral densities for $(1 - .5z)^{-1}$ as before, and also for $-(1 - .5z)$. Notice that the spectral density for $-(1-.5z)$ is identical to the spectral density for $(1-.5z)$.

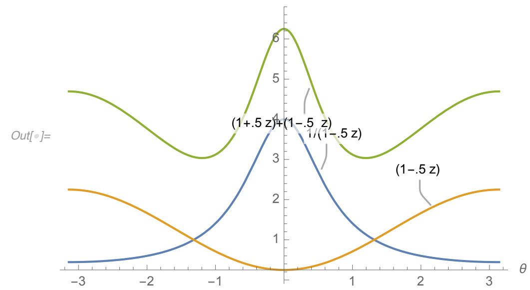

FIGURE 5. Spectral density of a sum.

However it is evident that the spectral density for the sum $(1-.5z)^{-1}-(1-.5z)$ is radically different from the spectral density for the sum $(1-.5z)^{-1}+(1-.5z)$ in Figure 5.

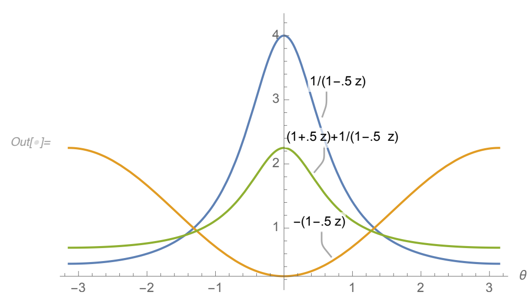

FIGURE 6. Spectral density of a sum.

There is however dimension in which spectral densities *can* be added. If we have two processes operating on two uncorrelated fundamentals then the spectral density of the sum *is* the sum of the spectral densities. This reflects the underlying property of variances: the variance of the sum of two independent random variables is the sum of the variances, and the spectral density at each frequency is a variance. This is illustrated in Figure 7.

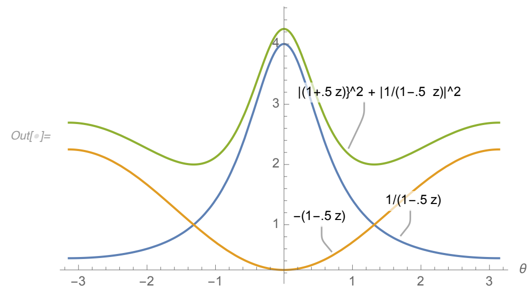

FIGURE 7. Spectral density of a sum of independent processes.

TABLE 1. Numerical algorithm tolerances

Spectral factorization	1×10^{-8}
Minimal realization	.0001
Balanced truncation	.1
Cauchy convergence	1×10^{-6}

TABLE 2. Base Parametrization

Discount factor η	Private variance σ_a^2	Public variance $\sigma_{\bar{a}}^2$	Noise variance σ_e^2	Private AR coefficient α	Public AR coefficient b
1	1.0	1.0	10.0	0.5	0.1

TABLE 3. Output filters

Private demand process $A(z)$	Public demand process $B(z)$
$\frac{2}{z-2}$	$\frac{10}{z-10}$
Direct intensity $\alpha(z)$	Direct intensity $\beta(z)$
$0.50 + \frac{0.018}{1.z-2.0} - \frac{0.02}{1.z-22.}$	$0.33 - \frac{0.10}{z-4.} + \frac{0.05}{z-2.19} - \frac{0.04}{z-110.}$
Total output on A	Total output on B
$0.05 + \frac{0.02}{z-2.04} - \frac{1.03}{z-2.}$	$0.07 - \frac{06.18}{z-10.0} - \frac{.35}{z-4.0} - \frac{0.0003}{z-2.04} - \frac{0.004}{z-2.00}$
Total output on e	Direct intensity δ
$0.002 - \frac{0.05}{z-2.04}$	$-0.001 + \frac{0.02}{1.z-2.0}$

Table 4. Base Parametrization

Discount factor η	Private variance σ_a^2	Public variance $\sigma_{\overline{a}}^2$	Noise variance σ_e^2	Private AR coefficient α	Public AR coefficient b
1	1.0	1.0	1.0	0.5	0.1

Table 5. Output filters

Private demand process $A(z)$	Public demand process $B(z)$
$\frac{2}{z-2}$	$\frac{10}{z-10}$
Direct intensity $\alpha(z)$	Direct intensity $\beta(z)$
$0.53 + \frac{0.20}{z-2.19} - \frac{0.17}{z-22.}$	$0.34 - \frac{0.07}{z-4.} + \frac{0.05}{z-2.19} - \frac{0.04}{z-110.}$
Total output on A	Total output on B
$0.05 - \frac{0.74}{z-2.19} - \frac{0.33}{z-2.} - \frac{0.005}{1.z-2.62}$	$0.07 - \frac{0.35}{z-4.} + \frac{0.002}{z-2.62} - \frac{0.042}{1.z-2.19} - \frac{6.38}{z-10.0}$
Total output on e	Direct intensity δ
$0.009 - \frac{0.56}{z-2.62}$	$-0.005 + \frac{0.28}{1.z-2.05}$

Table 6. Base Parametrization

Discount factor η	Private variance σ_a^2	Public variance $\sigma_{\overline{a}}^2$	Noise variance σ_e^2	Private AR coefficient α	Public AR coefficient b
1	1.0	1.0	1.0	0.1	0.5

Table 7. Output filters

Private demand process $A(z)$	Public demand process $B(z)$
$\frac{10}{z-10}$	$\frac{2}{z-2}$
Direct intensity $\alpha(z)$	Direct intensity $\beta(z)$
$0.52 + \frac{0.78}{z-10.58} - \frac{0.02}{1.z-110.}$	$0.33 + \frac{0.30}{z-10.5822} - \frac{0.12}{z-4.0} - \frac{0.006}{z-22.}$
Total output on A	Total output on B
$0.05 - \frac{4.43}{z-10.58} - \frac{0.72}{1.z-10.} - \frac{0.08}{z-12.17}$	$0.07 + \frac{0.03}{z-10.58} + \frac{0.16}{1.z-4.} - \frac{1.6}{z-2.} - \frac{0.003}{1.z-12.17}$
Total output on e	Direct intensity δ
$0.02 - \frac{2.10}{1.z-12.17}$	$-0.008 + \frac{1.08}{z-10.04}$

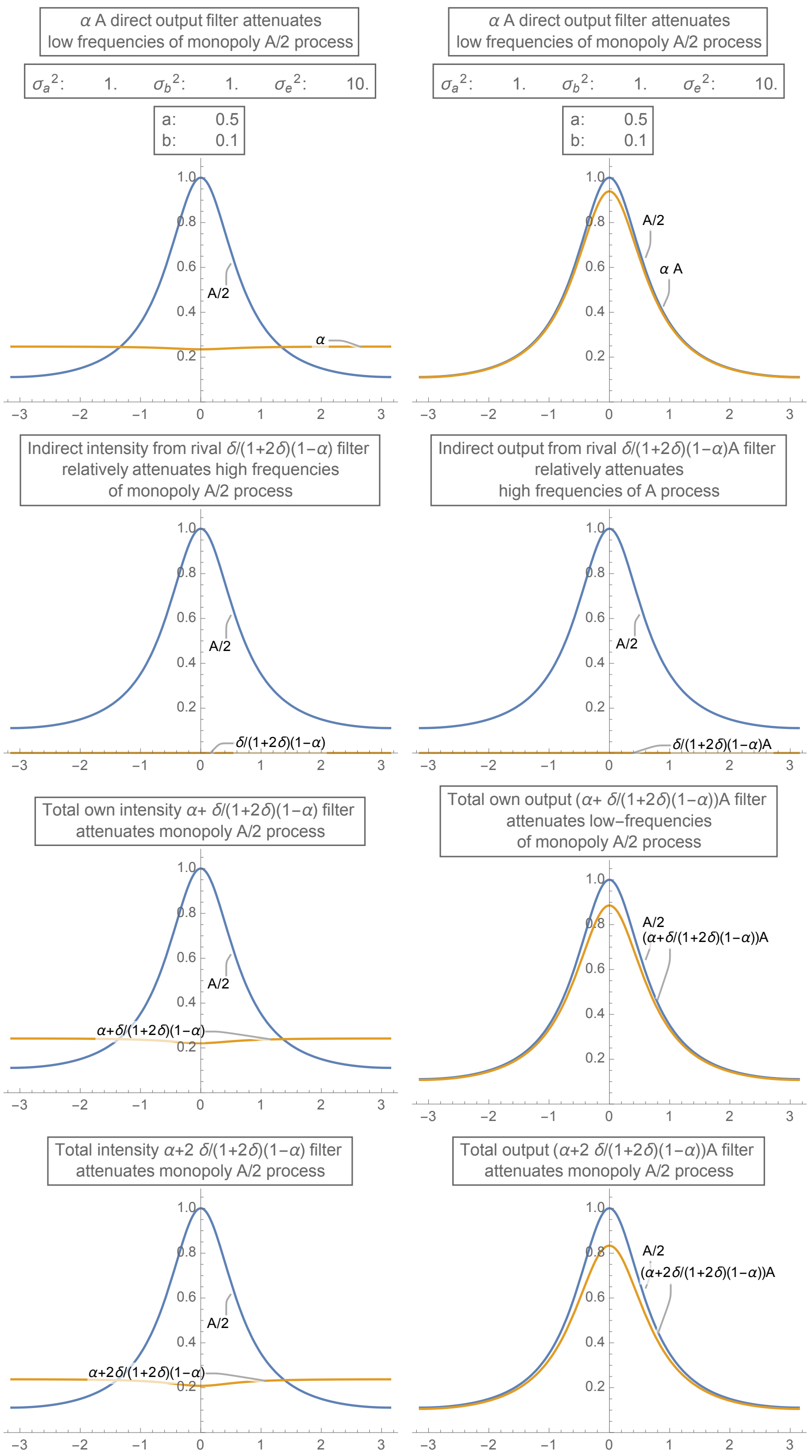

FIGURE 8. Spectral densities for private shock process $A(L)a_{it}$

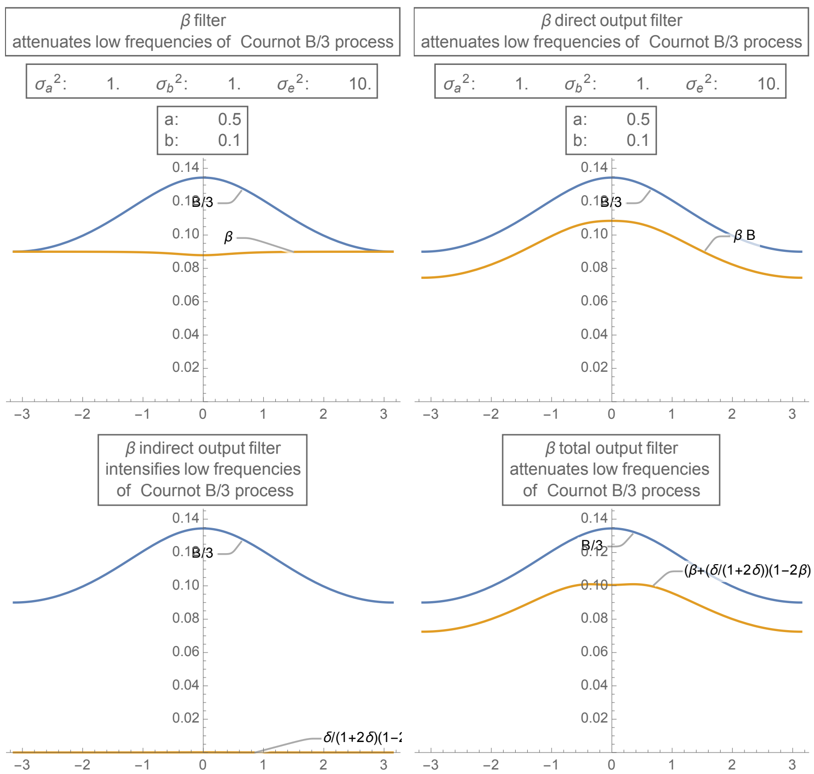

FIGURE 9. Spectral densities for public shock process $B(L)\overline{a}_t$

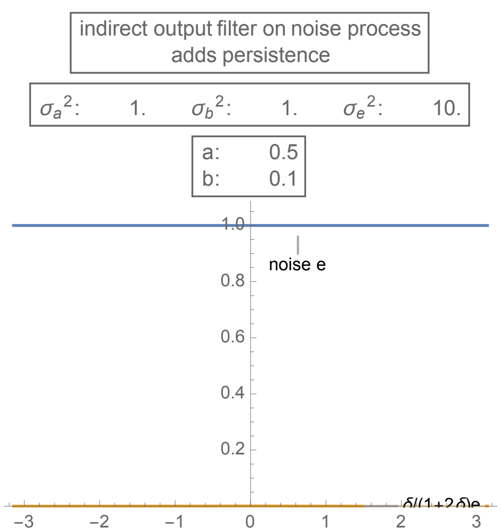

FIGURE 10. Spectral densities for noise process e_t

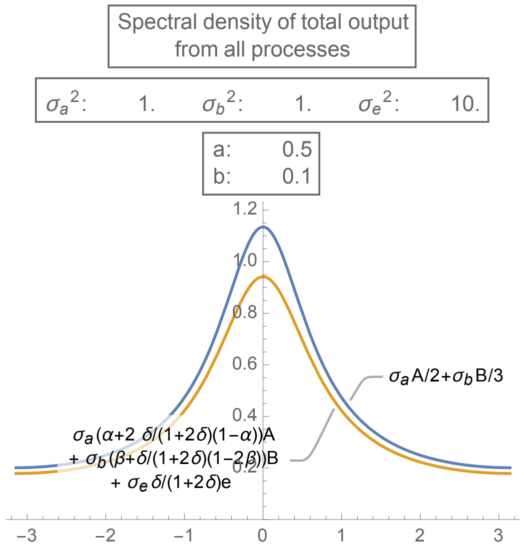

FIGURE 11. Spectral densities for total output processes on $A(L)a_{it}$, $B(L)\overline{a}_t$, and e_t

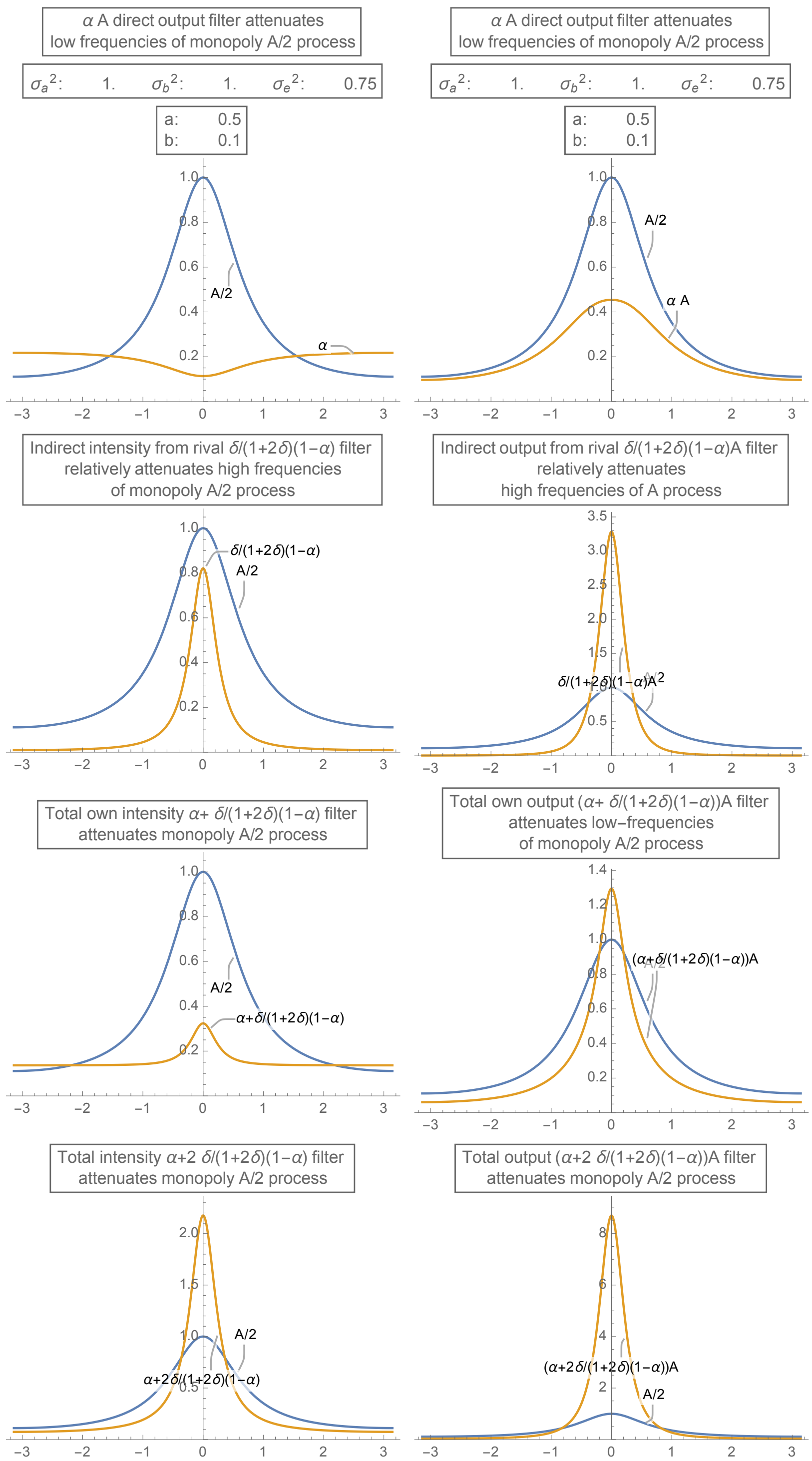

FIGURE 12. Spectral densities for private shock process $A(L)a_{it}$

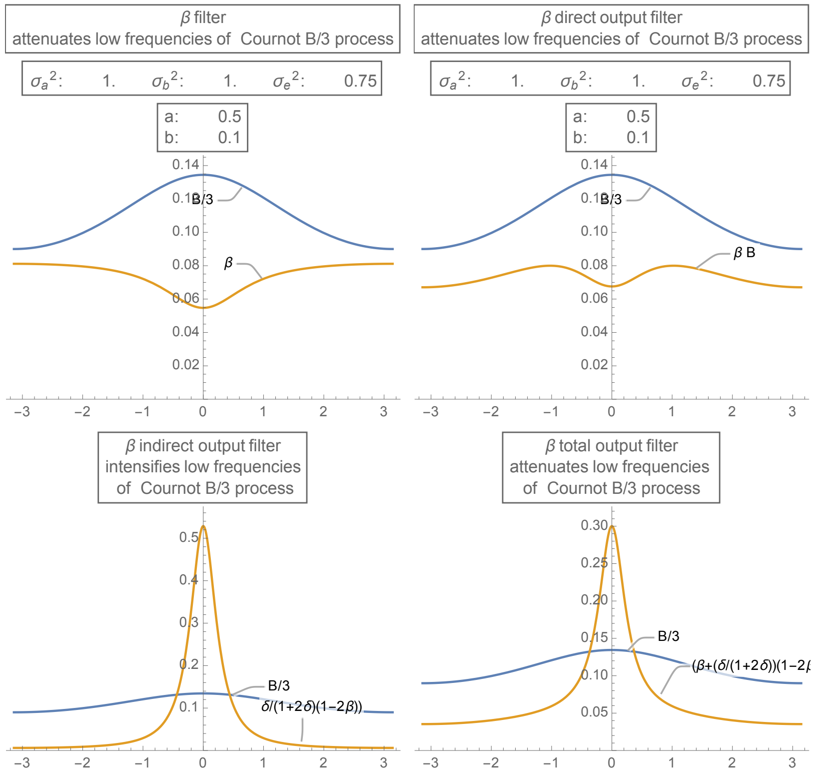

FIGURE 13. Spectral densities for public shock process $B(L)\overline{a}_t$

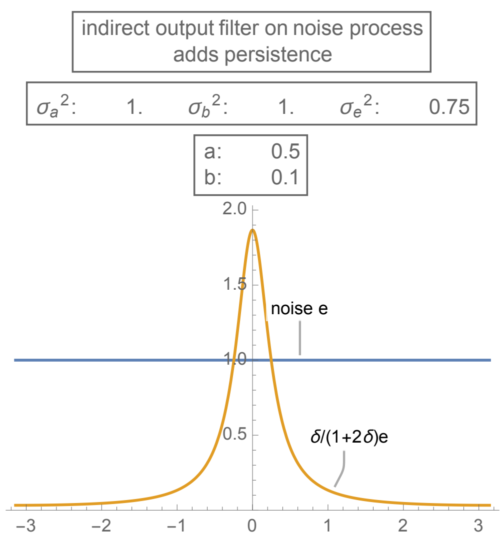

FIGURE 14. Spectral densities for noise process e_t

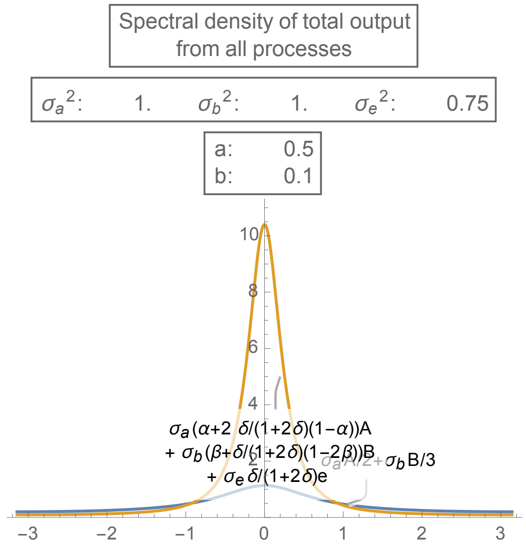

FIGURE 15. Spectral densities for total output processes on $A(L)a_{it}$, $B(L)\overline{a}_t$, and e_t

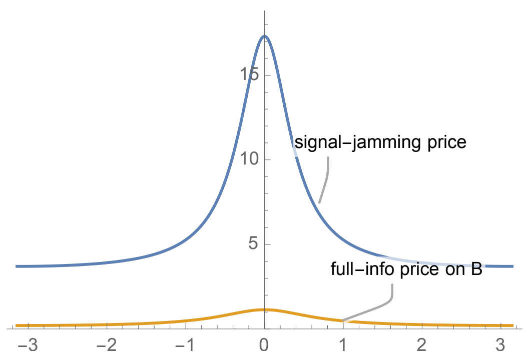

FIGURE 16. Spectral density for price process

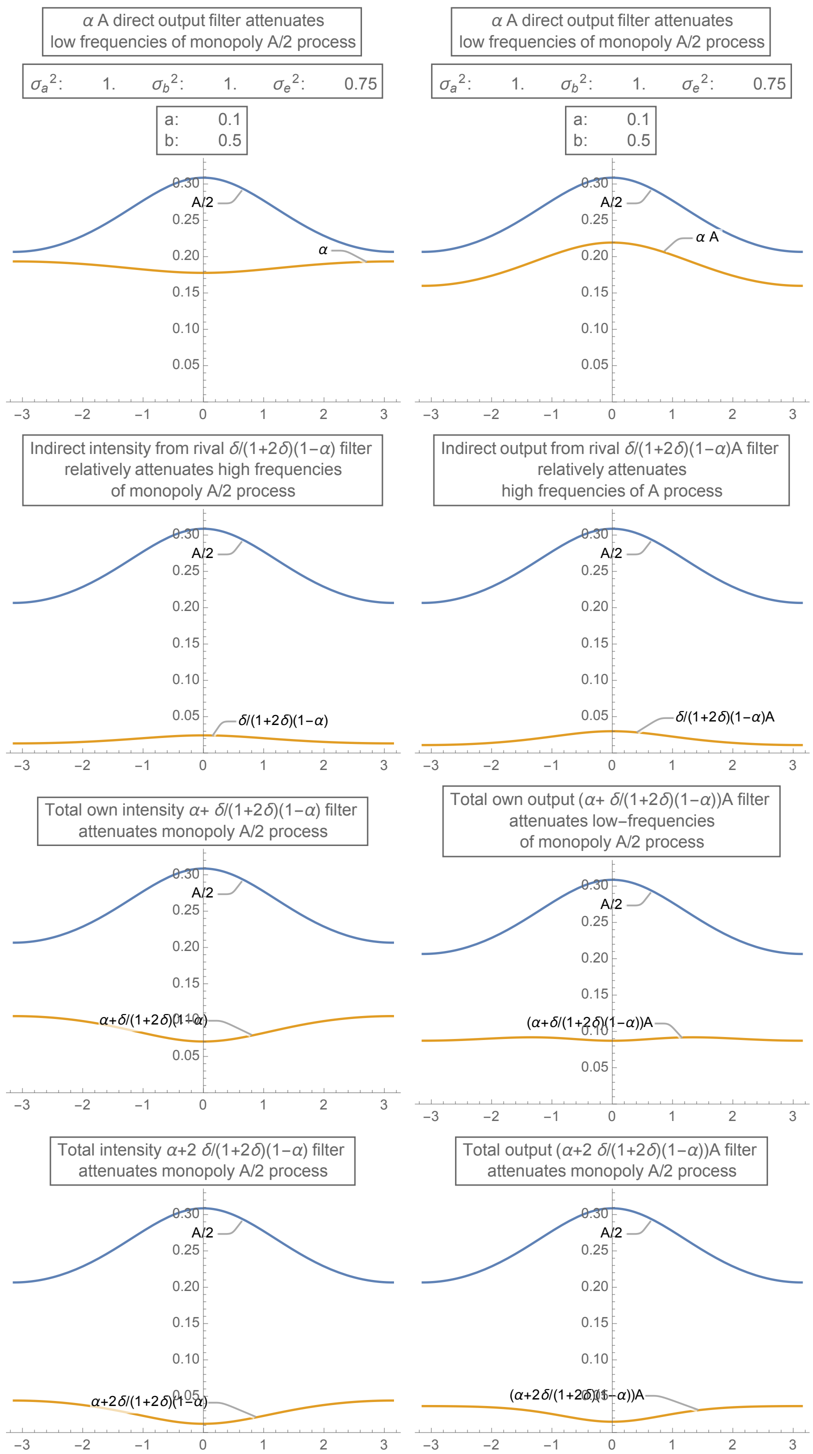

FIGURE 17. Spectral densities for private shock process $A(L)a_{it}$

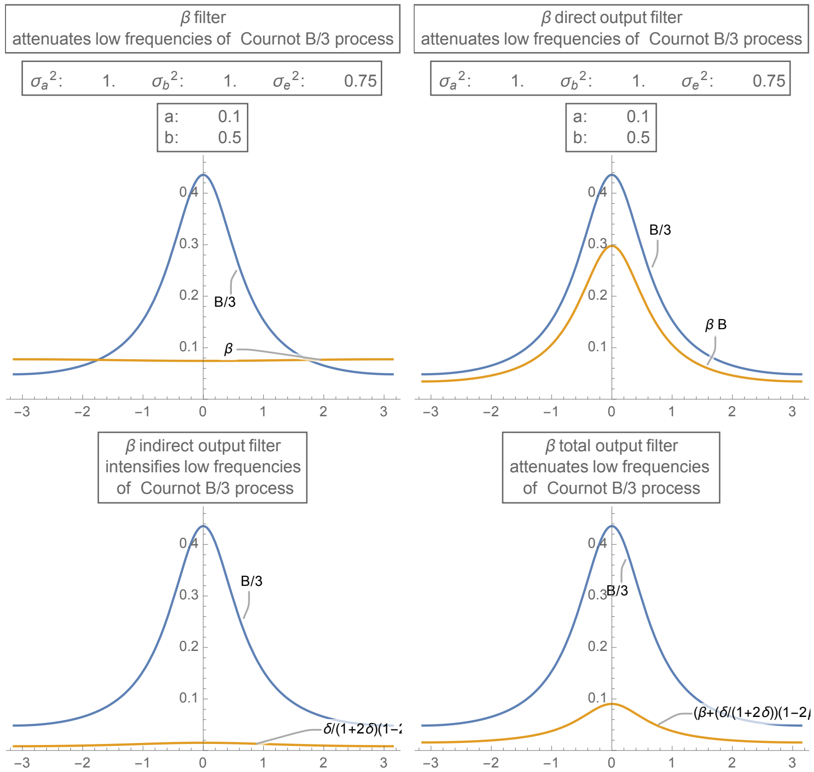

FIGURE 18. Spectral densities for public shock process $B(L)\overline{a}_t$

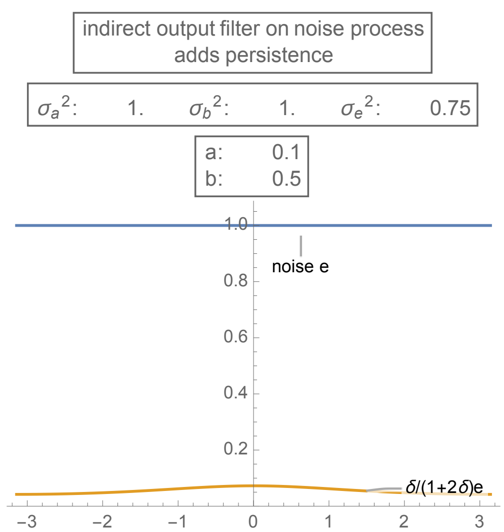

FIGURE 19. Spectral densities for noise process e_t

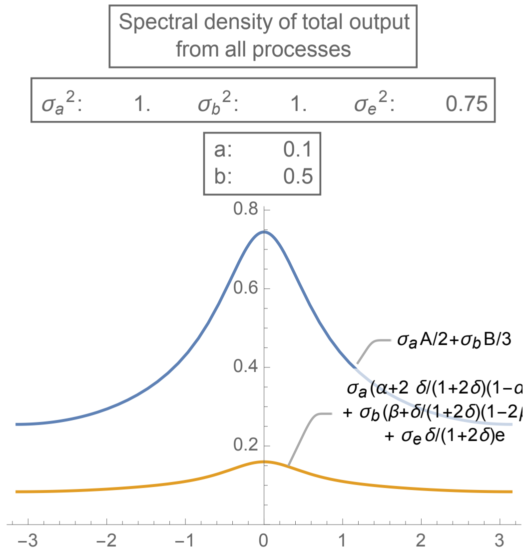

FIGURE 20. Spectral densities for total output processes on $A(L)a_{it}$, $B(L)\overline{a}_t$, and e_t