Games and Economic Behavior 142 (2023) 896–930

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

Signal-jamming in the frequency domain

B. Taub [1]

University of Glasgow, United Kingdom

ARTICLE INFO

Keywords:
Dynamic games
Signal jamming
Strategic information
Frequency-domain methods

ABSTRACT

I examine strategic behavior 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 elements of output over their full-information value, but also to *reduce* the persistence of price over its full-information value; this latter effect is "inconspicuousness."

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 and Urbano (1993) 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.[2] 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 character-

E-mail address: bart.taub@glasgow.ac.uk.

[1] 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.

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

https://doi.org/10.1016/j.geb.2023.10.006
Received 14 April 2022

Available online 21 October 2023

izing 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 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 and price 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.

1.1. Related literature

The analysis of oligopolistic competition in supply schedules with demand uncertainty dates back to Klemperer and Meyer (1989). 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, 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 (2015) analyze the static counterpart to the dynamic model in this paper. That model is related to Vives (2011). In this 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 its rival's cost signal.[3]

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 (1985), Aghion et al. (1991, 1993), Mirman and Urbano (1993), Caminal and Vives (2017), Harrington (1986a), Alepuz and Urbano (2005)). 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.[4]

Keller and Rady (1999) 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.[5] 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 (2008) 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 (1984), Abreu et al. (1986), Sannikov (2007), Hackbarth and Taub (2022)). 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 (2008) 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.

The model shares similarities of information and equilibrium structure with the financial speculation models descended from the model of Kyle (1989) and Kyle (1985). 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 Kyle (1989), 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 simultaneously 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 et al. (2017). 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.

[3] Bergemann et al. (2015) also analyze a static model with learning in which agents learn from private signals and prices.

[4] 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 (1986a,b), Caminal (2017), Bagwell and Ramey (1991), or Mailath (1989).

[5] Foundational papers on learning and experimentation by a monopolist include McLennan (1984), Aghion et al. (1991), Harrington (1995), Rustichini and Wolinsky (1995) and Keller and Rady (1999).

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 Kyle model this is achieved by structuring trades as functions of the forecast error of the uninformed market makers; here, as was demonstrated in the static model of Bernhardt and Taub (2015), firms structure output based on the rival's forecast error on the privately observed shocks. Consequently, even though demand shocks are highly serially correlated, price is not serially correlated at all. This is in stark contrast to the behavior of a full information model, in which output and price will have identical serial correlation.

1.2. 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.

1.2.1. Frequency domain applications in the literature

Hansen and Sargent (1980) 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 (1985), building on the work of Davenport and Root (1958), 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 (2000) 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 et al. (2014) 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 (2021) 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 (1989) study a model with small, risk-averse investors who behave competitively, showing there is no finite representation of the equilibrium. Huo and Takayama (2023) 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 (2008) who also demonstrate an infinite regress result.

Nimark (2017) 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 (2008) and Bernhardt et al. (2010) carry out an analogous iteration in the frequency domain, leading to a contraction property.

1.3. Plan of the paper

In Section 2, I set up the dynamic model and provide the solutions for equilibrium firm behavior. I then characterise the dynamic behavior of output and prices in the dynamic setting in Section 3. In Section 4, I illustrate the characterisations with numerical simulations for three canonical examples.

Following the conclusion there are five appendices. Two of these appendices are pedagogical in nature: Appendix A describes the frequency-domain methods used here in greater detail, and Appendix E describes the state-space numerical methods that must be used to simulate the model. The other appendices, Appendices B-D, 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 C.

2. The dynamic model

The dynamic model builds on the static model of Bernhardt and Taub (2015): there are two firms facing a demand curve for a homogeneous good that has a fixed and unitary slope, but a stochastic intercept. The intercept shock process has three independent elements: a common shock process $B(L)\overline{a}_t$ that is observable by both firms, a shock process $A_1(L)a_{1_t}$ that is observed privately by firm 1 but not by firm 2, and a shock process $A_2(L)a_{2_t}$ that is observed privately by firm 2 but not by firm 1. Each firm also observes a noisy signal of the price confounded by a noise shock e_t common to both firms. The underlying fundamental shocks $\{a_{it}, \overline{a}_t, e_t\}$, $t \in (\dots, -1, 0, 1, \dots)$, $i \in \{1, 2\}$, are mutually independent, serially-uncorrelated, zero-mean Gaussian processes, with variances σ_a^2, $\sigma_{\overline{a}}^2$ and σ_e^2, respectively.

Here, L denotes the lag operator, i.e., $Lx_t = x_{t-1}$, and the functions $A_i(\cdot)$, $B(\cdot)$ and so on denote linear functions of the lag operator, $A_i(L)a_{it} = \sum_{j=0}^{\infty} c_j a_{i,t-j}$, where c_j is the linear weight. It will be especially convenient to focus on processes that take first-order autoregressive form, that is, $A_i(L)a_{it} = \sum_{j=0}^{\infty} \rho^j a_{i,t-j}$, so that $A_i(L) = \frac{1}{1-\rho L}$, as the parameter ρ expresses the *persistence* of the process: a higher ρ results in a higher degree of persistence. Another way to refer to this persistence is to identify it with the *autoregressive structure* of the process, as the process $A_i(L)a_{it} = \frac{1}{1-\rho L}a_{it}$ is an autoregression (or AR process) whose characteristics—the persistence—are entirely determined by ρ.

The frequency-domain approach recognizes that the fundamental characteristics of a process like $A_i(L)a_{it}$ are determined by the function $A_i(L)$ alone, and maps that function into the set of functions of a complex variable. One is concerned with points in the complex plane at which such functions take on the value zero or infinity. These points are referred to as either *zeroes* or *poles*, respectively. From this point of view, $\frac{1}{\rho}$ is a *pole*. The inverse of the pole thus is simply the persistence, so characterising the pole is equivalent to characterising the persistence, and in this paper these terms will be used interchangeably.

Firms have a common discount factor η, and the exogenous functions of the lag operator such as $\frac{1}{1-\rho L}$ are assumed to satisfy $|\rho| < \eta^{-1/2}$ 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.

In addition to observing the publicly observable demand shock $B(L)\overline{a}_t$ and their privately observable shock $A_i(L)a_{it}$, the firms simultaneously observe a noisy signal of the price, $p_t + e_t$, with the noise shock e_t common to both firms, which they use to determine their output, whilst at the same time influencing the price with their output. 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],[7]

The firms compete in supply schedules. Denoting the vector of driving processes as $X_t \equiv \{A_1(L)a_{1t}, A_2(L)a_{2t}, B(L)\overline{a}_t, e_t\}$, 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:

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

and

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

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

I solve for an equilibrium in which the supply functions are linear and stationary,[8] taking the form

$$Q^i(p^t + e^t, X_i^t) = \alpha_i(L)A_i(L)a_{it} + \beta_i(L)B(L)\overline{a}_t + \delta_i(L)(p_t + e_t) \tag{3}$$

Definition 1. 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 (3) 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 (1)- (2) for all $t = \ldots, -1, 0, 1, \ldots$.

2.1. Solution procedure

To find the linear equilibrium I first 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 optimize over the supply function itself,[9] which in this linear setting translates into optimizing over linear functions of the histories of observed public and private shock realizations and of the history of the price signal.[10] I then translate the optimisation problem to the frequency domain and solve.

To begin, I verify the linearity of the output functions.

[6] Thus, as discussed in the introduction, the information structure is similar to that in the papers of Kyle (1989) and Bonatti et al. (2017).

[7] I assume that current and past realized profit do not result in any improvement in the signal of price. For example, ongoing inflation can add noise to the value of money in the calculation of profit.

[8] 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. For a related discussion see Foster and Viswanathan (1996), p. 1446.

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

[10] 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.

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

Proof. The proof of this lemma and other propositions is in Appendix B. □

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.[11] 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.

Reduced form. To transform the model to the frequency domain, the first step is to derive the reduced forms for output and price; first, apply Lemma 1 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}, then substitute the linear output function of firm i as a function of its information and the conjectured linear form of its rival (see Appendix B, equation (B.1)) into the price function to obtain its linear structure

$$p_t = \pi(q_{1t}, q_{2t}; X_t) = (1+\delta_1(L)+\delta_2(L))^{-1}\Big((1-\alpha_1(L))A_1(L)a_{1t} + (1-\alpha_2(L))A_2(L)a_{2t} + (1-\beta_1(L)-\beta_2(L))B(L)\overline{a}_t + e_t\Big). \tag{4}$$

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,

$$\begin{aligned} q_{it} = {} & \alpha_i(L)A_i(L)a_{it} + \beta_i(L)B(L)\overline{a}_t \\ & + \delta_i(L)(1+\delta_1(L)+\delta_2(L))^{-1}\Big((1-\alpha_1(L))A_1(L)a_{1t} + (1-\alpha_2(L))A_2(L)a_{2t} \\ & + (1-\beta_1(L)-\beta_2(L))B(L)\overline{a}_t + e_t\Big). \end{aligned} \tag{5}$$

2.2. Economic interpretation of the price and output functions

The output function, (5), has two components. The first component, $\alpha_i(L)A_i(L)a_{it} + \beta_i(L)B(L)\overline{a}_t$, expresses the direct effect of the intensity filters $\alpha_i(L)$ and $\beta_i(L)$ on each firm's privately observed shock $A_i(L)a_{it}$ and public shock $B(L)\overline{a}_t$ respectively. In a full-information model these intensities would simply be scalar constants, equal to the full-information monopoly value of $\frac{1}{2}$ and the duopoly value of $\frac{1}{3}$ respectively. Due to strategic effects however these intensity filters will not be scalars. As a result there will be strategic effects on the "direct" intensities α and β.

The second intensity component of the output function, $\delta_i(L)$, filters the noisy signal of price in equation (4). Because output is determined by this signal, when output is aggregated across firms to determine price, the effect of output on price is additionally influenced by the term $(1+\delta_1(L)+\delta_2(L))^{-1}$.

Examining equation (4), it is evident that there is feedback from both the direct intensities on private shocks, $\alpha_i(L)A_i(L)a_{it}$, and from the direct intensities on the public shock, $\beta_i(L)B(L)\overline{a}_t$, via the price, that is, there are indirect intensities on both public and private shocks. Each firm is able to extract these feedback effects for the shocks they directly observe, their own private shock and the public shock, with the net result that the rival firm's private output process, $\alpha_{-i}(L)A_{-i}(L)a_{i-t}$, is observed with noise via the price signal. The rival firm understands this and will therefore adjust its direct intensity on private information, $\alpha_{-i}(L)$, as well as via its intensity on publicly observable demand, $\beta_{-i}(L)$, which also affects the price, to influence this signal; this is signal jamming. Firm i understands this and will in turn adjust its own output intensity in response.

2.3. Mapping to the frequency domain

I next transform the objective to the frequency domain using these linear expressions. Each firm's objective is the sum of its 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 (4) and (5). For example, p_t contains the term

$$(1+\delta_1(L)+\delta_2(L))^{-1}(1-\alpha_i(L))A_i(L)a_{it}, \tag{6}$$

operating on a_{it} in equation (4), which is cross-multiplied by

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

$$\alpha_i(L)A_i(L)a_{it} + \delta_i(L)(1+\delta_1(L)+\delta_2(L))^{-1}(1-\alpha_i(L))A_i(L)a_{it} \tag{7}$$

from firm i's output q_{1t} in equation (5), also operating on a_{it}. The cross-products of these elements with all other terms are zero because the underlying stochastic processes are uncorrelated, that is, the expectation operator passes through the products, yielding terms like $E[a_{it}]E[\overline{a}_t]$, which are zero. 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

$$E\left[\left(H_j L^j a_{i,t+s}\right)\left(G_k L^k a_{i,t+\tau}\right)\middle|(a_{it}, \overline{a}_t, p_t + e_t)\right] \tag{8}$$

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

$$H_j G_k E\left[a_{i,t+s-j}a_{i,t+\tau-k}\middle|(a_{it}, \overline{a}_t, p_t + e_t)\right] \tag{9}$$

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., that is,

$$E\left[a_{i,t+s-j}a_{i,t+\tau-k}\right] = \begin{cases} 0, & s-j \neq \tau - k \\ \sigma_a^2, & s-j=\tau-k. \end{cases} \tag{10}$$

Importantly, the conditioning is irrelevant 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:

$$E\left[a_{i,t+s}a_{i,t+\tau}\right] \sim \sigma_a^2 \frac{1}{2\pi i}\oint z^s z^{-\tau}\frac{dz}{z} = \begin{cases} 0, & s \neq \tau \\ \sigma_a^2, & s=\tau, \end{cases} \tag{11}$$

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 (6) and (7) thus interact as an inner product, appearing as the convolution integral

$$\frac{1}{2\pi i}\oint D(1-\alpha_i)(\alpha_i^* + \delta_i^* D^*(1-\alpha_i^*))A_iA_i^*\sigma_a^2\frac{dz}{z}, \tag{12}$$

using the definition

$$D(z) \equiv (1+\delta_1(z)+\delta_2(z))^{-1}, \tag{13}$$

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

$$D^* \equiv D(\eta z^{-1}) \tag{14}$$

and so on for the other functions.[12]

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 four parts attached to the variance of each of the four mutually independent innovation processes a_{1t}, a_{2t}, $\overline{a}_t$ and e_t. Following the same procedure used to obtain (12), one obtains the frequency-domain version of firm i's objective:

$$\begin{aligned} \max_{\alpha_i,\beta_i,\delta_i} \frac{1}{2\pi i}\oint \Bigg(& D(1-\alpha_i)(\alpha_i^* + \delta_i^* D^*(1-\alpha_i^*))A_iA_i^*\sigma_a^2 \\ & + D(1-\alpha_{-i})\delta_i^* D^*(1-\alpha_{-i}^*)A_{-i}A_{-i}^*\sigma_a^2 \\ & + D(1-\beta_i-\beta_{-i})\left(\beta_i^* + \delta_i^* D^*(1-\beta_i^*-\beta_{-i}^*)\right)BB^*\sigma_{\overline{a}}^2 \\ & + (D-1)\delta_i^* D^* \sigma_e^2\Bigg)\frac{dz}{z}, \end{aligned} \tag{15}$$

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

[12] 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 (15) is over the intensity *functions* α_i, β_i, and δ_i—that is, it is a *variational* problem—taking as given the rival's intensity functions α_{-i}, β_{-i}, and δ_{-i}. This approach is equivalent to the time-domain approach in which output is chosen directly rather than via the intensity functions; the general equivalence of frequency-domain and time-domain optimisation is established and discussed in more detail in Appendix A.4.[13]

Variational derivatives. I set out the detailed variational derivatives—the Euler equations—of the frequency-domain objective (15) in Appendix B, equations (B.8) and (B.9). 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. 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

$$M y = Bx,$$

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^*, the conjugate transpose of F, corresponds to the (infeasible) weighting of *future* realizations; (ii) multiply both sides by the inverse of F^*; (iii) apply a projection to the resulting right-hand side—the $[\cdot]_+$ operator that eliminates terms that weight future histories[14]; and finally (iv) multiply both sides by the inverse of 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.[15]

2.4. Solution of a firm's variational optimization problem and equilibrium

It is useful to gather terms by implicitly defining two auxiliary functions, F and J, which solve

$$F^* F \equiv D(1 - \delta_1^* D^*) + D^*(1 - \delta_1 D) \tag{16}$$

$$J^* J \equiv (1 - \alpha_2^*)(1 - \alpha_2) A_2 A_2^* \sigma_a^2 + \sigma_e^2, \tag{17}$$

which have, as discussed above, factorizations F and J. The function F is the projection coefficient structure corresponding to the net information in the noisy price signal after a firm has extracted information from the price signal. The function J is the filter characterizing the information process from firm 1's observation of the price signals.

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

Proposition 1. *In a symmetric equilibrium firm i's filters on its direct information sources, a_i and $\bar{a}$ are given by*

$$\begin{aligned} \alpha &= 1 - F^{-1} A^{-1} \left[F^{*-1} D^* A \right]_+ \\ 1 - \beta &= \frac{1}{2} + \frac{1}{2} F^{-1} B^{-1} \left[F^{*-1} (1-\beta) D^* B \right]_+ \end{aligned} \tag{18}$$

Output weights on price signals satisfy the recursive system

$$J^* J = F^{-1} \left[F^{*-1} D^* A \right]_+ \left[F^{*-1} D^* A \right]_+^* F^{*-1} + \sigma_e^2 \tag{19}$$

$$D = \frac{1}{2} J^{-1} \left[J^{*-1} \sigma_e^2 \right]_+ + \frac{1}{2} J^{-1} \left[J \frac{D + D^*}{1 + D^*} \right]_+ . \tag{20}$$

Lemma 2. *If analytic functions $\{\alpha, \beta, \delta\}$ in $H^2[\eta]$ satisfy (18), (13), (19), (20), then the time-domain version of $Q^{i,t}$ defined in (3) 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. □

[13] 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 (B.7). 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.

[14] The projection or "annihilator" operator, $[\cdot]_+$, eliminates terms with negative powers of z from the Laurent expansion of a function: if $f(z) = \cdots + b_{-2} z^{-2} + b_{-1} z^{-1} + b_0 + b_1 z^1 + b_2 z^2 + \ldots$, then $[f]_+ = b_0 + b_1 z^1 + b_2 z^2 + \ldots$. 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.

[15] 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; this is explained in greater detail in Appendix A. 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 (1967).

Proposition 2. *A stationary linear equilibrium exists.*

The proof is in Appendix C. The fixed point argument uses the recursive system (19)–(20). Denoting the right-hand side of (20) by $S(D)$, write the recursion as

$$D = S(D),$$

defining a recursion in D (equation (19) 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.

3. Characterization of dynamic signal jamming and learning

In this section I characterise the behavior 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 behavior 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.

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 D.

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 3. *The equilibrium output intensity filters α_i and β_i are not scalar-valued, that is, output intensities are not just direct amplifications of the dynamic shock processes. This autoregressive structure of α_i and β_i therefore 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 using information from a noisy signal, 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.[16]

3.1. *Theoretical results on dynamic signal jamming*

I next 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 (18) and (19).) This in turn implies that δ and D

[16] That is, the filter is of the form $\sum_0^\infty \frac{c_i}{1-\rho_i z}$, $|\rho_0| > |\rho_1| > \ldots$.

will be scalar constants, and then finally that F is a scalar constant. (See equations (16) and (20).) Examining the formula for β in Proposition 1, 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

$$\begin{aligned}\beta &= F^{-1}B^{-1}\left[F^{*-1}(1-\beta)\left(D(1-\delta^* D^*)-D^*\delta D\right)B\right]_+ \\ &= F^{-1}B^{-1}F^{-1}(1-\beta)(D(1-\delta D)-D\delta D)\,B \\ &= F^{-1}F^{-1}(1-\beta)(D(1-\delta D)-D\delta D)\end{aligned} \tag{21}$$

which is a scalar constant. Thus, we have

Proposition 4. *Let the private demand shock process $A(L)a_{it}$ be i.i.d., that is, $A(L)$ is a scalar constant, and let the public demand shock process $B(L)\overline{a}_t$ be serially correlated. Then the intensity on the public shock $\beta(L)B(L)\overline{a}_t$ will be 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.

The second experiment reverses 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, 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 signal of the output process on the public shock via the price indistinguishable from the signal of the output on the private process.

This echoes the inconspicuousness findings of the literature on informed trading on private information in stock markets as in the model of Kyle (1989).

4. 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 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 publicly observed demand shock process is highly persistent whilst the privately observable shocks are not, echoing the setting in Proposition 4; in the final example I reverse the situation, echoing Proposition 5.

To characterise the results I 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. In addition to the pole characterization I illustrate how the firms dynamically re-weight the past realizations of the demand shocks as learning takes place, and illustrate this re-weighting with plots. Finally, I carry out a Monte Carlo experiment in which I generate a time series of demand shock innovations and measure the first-order serial correlation of the output and price constituents of the model, demonstrating how signal jamming modifies the serial correlation relative to a full-information model, and demonstrate the inconspicuousness result.

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 (20). These methods suppose that the stochastic processes can be represented as

$$x_t = Ax_{t-1} + Bu_t, \tag{22}$$

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 a first-order autoregressive process, that is, a process whose underlying filter is of the form $\frac{1}{1-\rho z}$.

To analyze the dynamics of output, one could find the fixed point of the recursion in (20), 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 3 establishes 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 (20).

To numerically approximate the equilibrium requires imposing an algorithm that trims quantitatively unimportant terms after each iteration of (20). 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 E.

4.1. Example 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 (therefore $\delta(z)$ will be approximately zero), 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 is $\alpha(z)A(z) = \frac{1}{2}A(z)$. 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 $\beta(z)B(z) = \frac{1}{3}B(z)$ for 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, and use the parameterization in Table 2.

The numerically calculated output intensity functions are presented in Table 3. Even though the private demand shock process is highly serially correlated (autoregressive parameter 0.5), with high noise, there is essentially no signal jamming: the intensity on A is close to the static monopoly value of $\frac{1}{2}$ with very little serial correlation, and the intensity on B is close to the static full-information duopoly value of $\frac{1}{3}$, again with very little serial correlation.

4.2. Example 2: persistent public demand shocks and low-persistence private shocks

For the second example I assume that the private shocks have very low persistence and that the public shocks have high innate persistence, and also that the noise variance is reduced by a factor of 10; as a result, there is potentially useful information in the price signal (see Table 4). Because the *private* shocks have low persistence, as established in Proposition 4 there is no need for firms to engage in *dynamic* signal jamming in the sense of altering the dynamic structure of the $B(L)\overline{a}_t$ process. Table 5 shows that there is very little deviation of the output intensity on the public demand shock from its full-information static Cournot value of $\frac{1}{3}$.

4.3. Example 3: persistent $A(L)a_t$ process, low-persistence $B(L)\overline{a}_t$ process

For the third example I examine the result of Proposition 5 numerically using the parameterization in Table 6. 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, just as in the first example, and again the noise variance is reduced by a factor of 10 relative to the first example; as a result, there is useful information in the price signal. (The serial correlation of the public shock process is not reduced to zero in order 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 results are displayed in Table 7. A key result is that the leading constant term in the intensity filter for the publicly observable shock $B(L)\overline{a}_t$ is .34, and the constant term for the privately observable shock $A(L)a_{it}$ is .53, essentially no different from the full-information duopoly and monopoly values of $\frac{1}{3}$ and $\frac{1}{2}$ respectively. If these were the only intensities the firms would simply amplify the public and private demand shocks, and there would be no difference between the autoregressive structure of the input processes and the output processes.

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, and the rival firm will wish to prevent this.

These effects are evident in the filters for output on the public shock in Table 7. The private direct intensity filter $\alpha(z)$ has main pole 2.19 expressing high persistence; the pole of the fundamental; the direct intensity $\beta(z)$ for the public shock also has a nontrivial term with the same pole. Thus, the output on the public shock has persistent elements even though the underlying shock, $B(L)\overline{a}_t$, has low intrinsic persistence.

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, even though the noise process e_t is itself serially uncorrelated.

4.4. Learning and signal jamming effects

As the firms observe the noisy price signal over time, they are able to learn about the rival's past realizations of its private demand shock process. At longer lags the precision with which a firm sees its rival's private demand shock innovation becomes sharper and sharper. The learned part of the shock is then equivalent to public information, and so the firm then signal jams on it.

To highlight this effect I calculated the weights on the innovations for the private and public output processes for ten lags, but normalized them using the weights of the underlying demand shock process. Thus, for example, I calculated the weights on the direct intensity filter $\alpha(z)$ for the private shock, applied to the demand shock process filter $A(z)$, but then normalized by the inverses of the weights of $A(z)$. Thus, if the normalization is applied to $A(z)$, which has lag coefficients

$$1, \frac{1}{2}, \frac{1}{4}, \ldots$$

then the normalized lag coefficients are

$$1, 1, 1, \ldots$$

and a plot of the lag weights would result in a horizontal line. Thus all lag weights are made comparable with this normalization. Under the normalization, if an output process filter such as the direct output filter $\alpha(z)A(z)$ has less persistence than the fundamental demand shock process as expressed by $A(z)$, then the normalized weights will be decreasing, whereas if it has greater persistence then they will be increasing.

A key feature of classical signal jamming is that firms increase output on publicly observable demand in order to mask their output on privately observable demand. Operationally this translates into firms *reducing* their direct output intensity on public shocks so that the output signal in price *increases*; examining equation (4), the effect on price enters via the $(1 - \beta_1(L) - \beta_2(L))B(L)$ term, which increases as $\beta(L)$ decreases. Thus, a prediction of the model is that the direct intensity on public shocks will decrease as signal jamming increases.

In Example 3, the learned innovations result in highly persistent shocks, and the rival therefore has an incentive to magnify its signal jamming intensity on that shock. Thus, the indirect intensity on private shocks actually increases with longer lags, asymptotically approaching a number larger than the $\frac{2}{3}$ value that would hold under full information; see Fig. 1.

Meanwhile, the total intensity on the actual public shock shrinks with longer lags; this happens because the value of signal jamming at long lags shrinks because the shock is not persistent. Thus, there are two effects: the first effect increases the persistence of the output response on public shocks due to the desire to camouflage the public shock, as noted in Proposition 5; the second effect shrinks the intensity at long lags nevertheless (when measured using the *private* shock process decay rate). Thus, in the primary example, most of the signal jamming in the sense of increased intensity takes place on the learned private shock, not the public shock, with the effect working through the $\frac{\delta}{1+2\delta}$ coefficient (see equation (5)), which also takes on the persistence of the private shock—see the numerical estimate of δ in Table 7.

Expanding on this point, the signal jamming effect on the public shock takes place via the price component, that is, the indirect output on the public shock. As is evident from Fig. 2, the indirect output on the public shock, when normalized by the weights from the *private* shock process for the example with persistent private but non-persistent public demand shocks, the indirect component of output on public shocks closely matches the persistence of the private process: the normalized indirect output intensities are flat (even though their magnitude is small), indicating high persistence, whereas the direct intensity falls essentially to zero after the first lag, reflecting that the underlying public shocks are close to serially uncorrelated, and so the intensity is close to the pure static intensity, with no induced serial correlation. This reflects the desire to disguise output on the public process.

4.5. Inconspicuousness

The intuition so far is that the desire to disguise the output from the public process as if it were output on the private process adds persistence to the public output process. This can be verified via Monte Carlo simulations: in the main numerical experiment (Example 3, see Table 6), I generated appropriate time series examples with 1,000 pseudorandom innovations for each of the two private paths, the public demand shocks, and the noise paths. I then calculated the serial correlation via the first-order autoregressive coefficient for each of the constituent time series in the output and price equations (5) and (4), numerically approximating the intensity filters in Table 7, via a moving average approximation with a lag length of 10. The results are in Table 8.

The first observation is that the serial correlation of the indirect output on the private shock process is .78, much higher than the intrinsic serial correlation of .5 of the fundamental shock; this reflects signal jamming on learned information. Despite this magnified serial correlation of the indirect term, the direct output serial correlation, .41, is much lower than the intrinsic serial correlation of the shocks; this reflects the stepping back of the firm on direct output as the rival learns and signal jamming takes over. The net result is that total output has serial correlation basically identical to the serial correlation of the underlying demand shock process.

The output on the public shocks is similar: the serial correlation of the total output on the public shock process is .14, much higher than the intrinsic serial correlation of .1 of the fundamental shock; this reflects the signal jamming that is taking place via the indirect output on the public shock, which has a serial correlation of .55—much closer to the serial correlation of the *private* shock process. However, the noisy signal comes through price, and so it is important to assess the serial correlation of price; this correlation is only .21. For comparison, if the firms operated without price feedback they would choose static output intensities of $\frac{1}{2}$ on the private shocks and $\frac{1}{3}$ on the public shocks, and there would be no noise component of output or price. In this case, with the parameterization of the main example (Table 6), the serial correlation from Monte Carlo simulations is .48, reflecting the influence of the serially correlated private demand shocks dominating the less serially correlated public shocks. Thus, in the noisy price signal model the serial correlation of price is *lower* than in the no-feedback case.

The explanation is related to a similar phenomenon in the Kyle (1985) model of financial market microstructure. In that model, the informed trader trades in such a way that his trades are driven by the fundamental value of the underlying asset, but in such a way that the market makers cannot observe it—it is based on the forecast error of the market makers. As a result, the information in the price signal observed by the market makers cannot be distinguished from the noise stemming from the trades of uninformed traders. Similarly here, the firms attempt to structure the price process so that it has no characteristics that would allow the firms to invert the signal to impute the rival's private process. Thus, the price process resembles only the public process, which has no dynamic structure that is usable to infer the rival's private process. This point was emphasized in Bernhardt and Taub (2015): each firm bases its output on its privately observed process on the forecast error of the rival firm.[17]

The firms do not see price directly, but rather a noisy signal of price; each firm knows the effect of its own output on price and also the effect of the public output on price, so these can be netted out. In the Monte Carlo simulations, the serial correlation of the net information in the noisy price signal is .04, that is, very close to the zero correlation of pure noise.

5. Conclusion

There are three phenomena. The first is signal jamming, in which firms raise output keyed to publicly observable demand shocks in order to mask their privately observed demand shocks. This effect is expressed in the firms' indirect output intensities on the public demand shock, that is, the part of output that influences price: the firms do not just increase the output intensity on those shocks, they also increase the serial correlation of the output relative to the fundamental serial correlation of the demand shock itself.

The second phenomenon concerns learning. The firms also learn about their rival's private shocks via the noisy price history. The learned information is then public, and firms then signal jam on the learned output. Concomitantly, each firm cuts back output direct output on its own private demand shocks as the rival learns, and in the long run the signal jamming takes over completely.

Finally, these strategic interactions have a significant impact on the behavior of price: due to the desire of firms to hide their private information, that is to be inconspicuous, even though demand shocks are highly serially correlated, price is not serially correlated at all. This is in stark contrast to the behavior of a full information model, in which output and price will have identical serial correlation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Bart Taub reports financial support and administrative support were provided by National Research University Higher School of Economics. Bart Taub reports financial support and administrative support were provided by European University Institute.

Data availability

No data was used for the research described in the article.

Appendix A. Frequency-domain methods

This is an augmented version of a similar appendix that appeared in Seiler and Taub (2008), which in turn built on the appendix in Whiteman (1985).

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

$$a_t = \sum_{k=0}^{\infty} A_k e_{t-k}. \tag{A.1}$$

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, 1974, 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

[17] See equation (11) and the surrounding discussion in Bernhardt and Taub (2015).

[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.

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]

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 (1967) 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 (A.1); 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,

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

subject to the budget constraint,

$$c_t = y_t + r b_{t-1} - b_t \tag{A.3}$$

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 objective, leaving the modified problem,

$$\max_{\{b_t\}} -E \sum_{t=0}^{\infty} \beta^t (y_t + r b_{t-1} - b_t)^2, \tag{A.4}$$

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 $k > s$, $k > r$,

$$E_{t-k}\left[e_{t-r}e_{t-s}\right] = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s, \end{cases} \tag{A.5}$$

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

$$E_{t-k}\left[(L^r e_t)(L^s e_t)\right] = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s. \end{cases} \tag{A.6}$$

Notice that the "action" is in the exponents of the lag operators. From Cauchy's theorem (Conway, 1985), it is equivalent to write

[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.

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

[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.

$$\sigma_e^2 \frac{1}{2\pi i} \oint z^r z^{-s} \frac{dz}{z} = \begin{cases} 0, & r \neq s \\ \sigma_e^2, & r = s, \end{cases} \tag{A.7}$$

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 (A.6) and (A.7) is crucial but might not be particularly intuitive. To see the equivalence, begin by calculating the following integral:

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

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

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

(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 (A.6) and (A.7).

Whiteman (1985) 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

$$E_t\left[\sum_{s=1}^{\infty} \beta^s a_{t+s} b_{t+s}\right] = E_t\left[\sum_{s=1}^{\infty} \beta^s \left(\sum_{k=0}^{\infty} A_k e_{t+s-k}\right)\left(\sum_{k=0}^{\infty} B_k e_{t+s-k}\right)\right]. \tag{A.8}$$

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

$$\begin{aligned} &\beta[A_0 B_0 + \beta A_1 B_1 + \beta^2 A_2 B_2 + \dots] E_t[e_{t+1}^2] \\ &\qquad + \beta^2 [A_0 B_0 + \beta A_1 B_1 + \beta^2 A_2 B_2 + \dots] E_t[e_{t+2}^2] + \dots \\ &= \beta[A_0 B_0 + \beta A_1 B_1 + \beta^2 A_2 B_2 + \dots]\sigma_e^2 \end{aligned}$$

$$
\begin{aligned}
&+\beta^2[A_0B_0+\beta A_1B_1+\beta^2A_2B_2+\dots]\sigma_e^2+\dots \\
&=\frac{\beta\sigma_e^2}{1-\beta}\sum_{s=0}^{\infty}\beta^k A_k B_k=\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint A(z)B(\beta z^{-1})\frac{dz}{z}.
\end{aligned}
\tag{A.9}
$$

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 (A.8), the representation in (A.9) permits a direct variational approach. Equation (A.9) 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 (1980) 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 (A.4) 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 (A.9), the agent's objective can be written as

$$\max_{B(\cdot)}-\frac{\beta\sigma_e^2}{1-\beta}\frac{1}{2\pi i}\oint(A(z)-(1-rz)B(z))(A(\beta z^{-1})-(1-r\beta z^{-1})B(\beta z^{-1}))\frac{dz}{z}. \tag{A.10}$$

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

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 $B(\cdot)$:

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

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

$$0=-\frac{1}{2\pi i}\oint(A(z)-(1-rz)B(z))(1-r\beta z^{-1})\zeta(\beta z^{-1})\frac{dz}{z}, \tag{A.11}$$

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

---

[22] A similar variational approach in continuous time can be found in Davenport and Root (1958), p. 223.

[23] By reformulating the problem, the Szegö-Kolmogorov-Krein theorem (Hoffman, 1962, 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 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.

The integral in first-order condition (A.11) 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 (A.11) 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 (A.11) from the backward-looking components, so that we can then eliminate the non-analytic portion from our solution. Examining equation (A.11), note that by construction ζ is analytic, so that it can be represented as a power series,

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

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 (A.11), $(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z))$, 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 (A.11) can now be broken out of the integral and stated as follows:

$$(1 - r\beta z^{-1})(A(z) - (1 - rz)B(z)) = \sum_{-\infty}^{-1}, \tag{A.12}$$

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 et al., 1990).

The Wiener-Hopf equation (A.12) can be restated as:

$$(1 - r\beta z^{-1})(1 - rz)B(z) = (1 - r\beta z^{-1})A(z) + \sum_{-\infty}^{-1}. \tag{A.13}$$

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 $\{z \big| |z| \le \beta^{1/2}\}$)) in negative powers of z. Hence, we can divide through by this factor to rewrite the Wiener-Hopf equation as

$$\beta r^2(1-(\beta r)^{-1}z)B(z) = \frac{(1-r\beta z^{-1})}{1-(\beta r)^{-1}\beta z^{-1}}A(z) + \sum_{-\infty}^{-1}, \tag{A.14}$$

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 (A.14) is the product of analytic functions, applying the annihilator to (A.14) 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 (2008). □

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.

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

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 (A.4),

$$\max_{\{b_t\}} -E\sum_{t=0}^{\infty}\beta^t(y_t + rb_{t-1} - b_t)^2, \tag{A.15}$$

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

$$b_t = B(L)e_t, \tag{A.16}$$

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

$$E_t[(1 - rL)(1 - r\beta L^{-1})B(L)e_t] = E_t[(1 - r\beta L^{-1})A(L)e_t]. \tag{A.17}$$

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 (A.17) 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 (A.17) 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 (A.13), 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 (A.13), this equation has a solution, validating the conjecture expressed in equation (A.16) 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 (A.10); it is helpful to express it in the equivalent form

$$\max_{\{B(\cdot)\}} -\|A - RB\|_2^2 \tag{A.18}$$

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. Dynamic model basic derivations and proofs

B.1. Proof of Lemma 1

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

$$q_{-it} = Q^{-i}(p^t + e_t; X_i^t) = \alpha_{-i}(L)A_{-i}(L)a_{-it} + \beta_{-i}(L)B(L)\overline{a}_t + \delta_{-i}(L)(p_t + e_t). \tag{B.1}$$

Substituting into price in (2) yields

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

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

$$\max_{q_{it}} E\left[\sum_{s=0}^{\infty}\eta^{t+s}\Big((1+\delta_{-i}(L))^{-1}\Big(A_i(L)a_{i,t+s} + (1-\alpha_{-i}(L))A_{-i}(L)a_{-i,t+s}\right.$$
$$\left. +(1-\beta_{-i}(L))B(L)\overline{a}_{t+s} - \delta_{-i}(L)e_{t+s}^P - q_{i,t+s}\Big)\Big)\, q_{i,t+s}\middle|\, p^t + e^t, X_i^t\right], \tag{B.3}$$

using the structure of price from equation (B.2) (but leaving the price function in the conditioning information abstract to conserve notation).

Define

$$\kappa(L) \equiv (1+\delta_{-i}(L))^{-1} = \sum_{s=0}^{\infty}\kappa_s L^s, \tag{B.4}$$

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

$$x_{it} \equiv A_i(L)a_{it} + (1-\alpha_{-i}(L))A_{-i}(L)a_{-it} + (1-\beta_{-i}(L))B(L)\overline{a}_t - \delta_{-i}(L)e_t. \tag{B.5}$$

Then firm i's objective can be written compactly as

$$\max_{q_{it}} E\left[\sum_{s=0}^{\infty}\eta^{t+s}\left(\kappa(L)(x_{i,t+s} - q_{i,t+s})\right) q_{i,t+s}\middle|\, p^t + e^t, X_i^t\right]. \tag{B.6}$$

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

$$0 = E\left[\kappa(L)(x_{it} - q_{it}) - \sum_{s=0}^{\infty}\eta^s\kappa^s q_{i,t+s}\middle|\, p^t + e^t, X_i^t\right]. \tag{B.7}$$

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 (B.7) 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. □

Proof of Proposition 1. The variational first-order conditions,[25] for α_1 and β_1 with respect to the objective (15) are:

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

[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.

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

$$\beta:\ \beta_1\Big(D(1-\delta_1^* D^*)+D^*(1-\delta_1 D)\Big)B^* B\sigma_a^2=(1-\beta_2)\left(D(1-\delta_1^* D^*)-D^*\delta_1 D\right)B^* B\sigma_a^2+\sum_{-\infty}^{-1}$$

Rewrite these Wiener-Hopf equations as

$$\alpha_1 F^* F A^* A\sigma_a^2=(D-(\delta_1+\delta_1^*)D^* D)A^* A\sigma_a^2+\sum_{-\infty}^{-1} \tag{B.8}$$

$$\beta_1 F^* F B^* B\sigma_a^2=(1-\beta_2)\left(D(1-\delta_1^* D^*)-D^*\delta_1 D\right)B^* B\sigma_a^2+\sum_{-\infty}^{-1} \tag{B.9}$$

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 1.

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 (20).

The δ_1 Wiener-Hopf equation is

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

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

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

Define H by

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

and rewrite the Wiener-Hopf equation as

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

with solution

$$\delta_1=H^{-1}D^{-1}\left[DH(1+\delta_2^*)-H^{*-1}\left(\sigma_e^2(1+\delta_2^*)+\alpha_1(1-\alpha_1^*)A_1A_1^*\sigma_a^2+\beta_1(1-\beta_1^*-\beta_2^*)BB^*\sigma_a^2\right)\right]_+ \tag{B.10}$$

To isolate the α and β equations, begin by rearranging the Wiener-Hopf equation for α, equation (B.8) 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:

$$(1-\alpha_1)D^* D\Big(2+\delta_1^* + \delta_1)\Big)A^* A\sigma_a^2 = D^* A^* A\sigma_a^2 + \sum_{-\infty}^{-1}. \tag{B.11}$$

Dividing out D^* and grouping terms yields

$$D\Big(1+\delta_2^* - \delta_1 - \alpha_1(2+\delta_2+\delta_2^*))\Big)A^* A\sigma_a^2 = \sum_{-\infty}^{-1}. \tag{B.12}$$

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

$$\big(D^* D(1-\alpha_1^*)(1-\alpha_1)(\delta_1 - \delta_2^* - 1) + D^*(1-\alpha_1^*)\alpha_1\big)\ A_1^* A_1\sigma_a^2.$$

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 (B.8). Thus, these terms all drop out of the δ_1 Wiener-Hopf equation (B.10).

The β_1 Wiener-Hopf equation is

$$\Big((1-\beta_1-\beta_2)D^* D(2+\delta_2+\delta_2^*) - D^*(1-\beta_2)\Big)B^* B\sigma_a^2 = \sum_{-\infty}^{-1}. \tag{B.13}$$

The β terms from the δ_1 Wiener-Hopf equation are

$$\Big(D^* D(1-\beta_1^* - \beta_2^*)(1-\beta_1-\beta_2)(\delta_1 - \delta_2^* - 1) + D^*\beta_1(1-\beta_1^*-\beta_2^*)\Big)B^* B\sigma_a^2.$$

Consolidating terms yields

$$D^*(1-\beta_1^*-\beta_2^*)\Big(D(1-\beta_1-\beta_2)(\delta_1-\delta_2^*-1) + \beta_1.\Big)B^* B\sigma_a^2.$$

Adding and subtracting $1-\beta_2$ yields

$$D^*(1-\beta_1^*-\beta_2^*)\Big(D(1-\beta_1-\beta_2)(\delta_1-\delta_2^*-1) - (1-\beta_1-\beta_2) + (1-\beta_2)\Big)B^* B\sigma_a^2.$$

Consolidating yields

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

with the last equality following from (B.13). Thus, the β elements also drop out of the δ_1 equation (B.10).

With the extraneous terms eliminated, the δ_1 Wiener-Hopf equation (B.10) reduces to

$$D^* D\Big((1-\alpha_2^*)(1-\alpha_2)A_2 A_2^*\sigma_a^2 + \sigma_e^2\Big)(1+\delta_2^* - \delta_1) = D^*\sigma_e^2(1+\delta_2^*) + \sum_{-\infty}^{-1}. \tag{B.14}$$

Next, substitute the definition of J from (17) in the main text, into equation (B.14) to obtain

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

Grouping the terms above yields:

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

Solving yields

$$\delta_1 = D^{-1}J^{-1}\left[(DJ - J^{*-1}\sigma_e^2)(1+\delta_2^*)\right]_+. \tag{B.15}$$

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

$$\delta=-\frac{1}{2}+\frac{D^{-1}}{2\sigma_e^2}\left[DJ^*J(1-\frac{\delta}{1+\delta^*})\right]_+ . \tag{B.16}$$

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

B.2. Derivation of the recursion in D

Manipulating the definition of D in equation (B.16) 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

$$\sigma_e^2=\left[DJ^*J\left(1-\frac{\delta}{1+\delta^*}\right)\right]_+ . \tag{B.17}$$

Next undo the annihilator operator and write

$$(1+\delta^*)\sigma_e^2+\sum_{-\infty}^{-1}=J^*J\frac{1+\delta^*-\delta}{1+2\delta}. \tag{B.18}$$

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 (B.18), to obtain

$$J^*J\left(1+\frac{D^{*-1}-1-(D^{-1}-1)}{2}\right)D=\frac{1+D^*}{2D^*}\sigma_e^2+\sum_{-\infty}^{-1}, \tag{B.19}$$

which reduces to

$$J^*J\left(D^*D+\frac{D}{2}-\frac{D^*}{2}\right)=\frac{1+D^*}{2}\sigma_e^2+\sum_{-\infty}^{-1}. \tag{B.20}$$

Starting with equation (B.20) we derive a new recursion in D. We first multiply equation (B.20) by 2:

$$J^*J\left(2D^*D+D-D^*\right)=(1+D^*)\sigma_e^2+\sum_{-\infty}^{-1}. \tag{B.21}$$

Next, add and subtract D:

$$J^*J\left(2D^*D+2D-(D^*+D)\right)=(1+D^*)\sigma_e^2+\sum_{-\infty}^{-1} \tag{B.22}$$

and then rearrange to obtain

$$2J^*JD\left(1+D^*\right)=(1+D^*)\sigma_e^2+J^*J(D^*+D)+\sum_{-\infty}^{-1}. \tag{B.23}$$

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

$$J^*JD = \frac{1}{2}\sigma_e^2 + \frac{1}{2}J^*J\frac{D^*+D}{1+D^*} + \sum_{-\infty}^{-1}. \tag{B.24}$$

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

$$\boxed{D = \frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+ + \frac{1}{2}J^{-1}\left[J\frac{D^*+D}{1+D^*}\right]_+.} \tag{B.25}$$

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 (16) and (17) to obtain

$$F^*F \equiv D(1-\delta^*D^*) + D^*(1-\delta D) \tag{B.26}$$

$$J^*J \equiv (1-\alpha^*)(1-\alpha)AA^*\sigma_a^2 + \gamma^*\gamma CC^*\sigma_c^2 + \sigma_e^2. \tag{B.27}$$

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

$$J^*J \equiv (1-\alpha^*)(1-\alpha)AA^*\sigma_a^2 + \sigma_e^2. \tag{B.28}$$

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 (B.26) 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

$$(1-\alpha) = cF^{-1}, \tag{B.29}$$

where c is a constant. Thus,

$$(1-\alpha)AA^*(1-\alpha^*) = F^{-1}\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^* F^{*-1}, \tag{B.30}$$

the left-hand side of which appears in equation (B.28).

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

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

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,

$$\boxed{F^*F = \left(D(1-\delta^*D^*) + D^*(1-\delta D)\right) = \frac{1}{2}(D^*+D+2D^*D).} \tag{B.31}$$

Thus,

$$\boxed{J^*J = F^{-1}\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^* F^{*-1} + \sigma_e^2} \tag{B.32}$$

Equations (B.25), (B.31) and (B.32) 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 C. 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 (B.25), 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 (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 (B.25) is a contraction, just as in the scalar model. In a functional recursion such as (B.25), 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

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

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 and ρ are zero. Intuition about the contraction property can be gleaned by considering the ordinary derivative of a scalar version of (B.25). 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 (B.25): 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 (B.25). In a scalar setting, substituting from (B.31), equation (B.30) becomes

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

where we arbitrarily write the scalar value of A as $\frac{1}{1-a}$. The recursion equations (B.32) and (B.25) then become a difference equation system,

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

and

$$D_{n+1} = \frac{1}{2}\frac{1}{J_n^2}\sigma_e^2 + \frac{D_n}{1+D_n}. \tag{C.3}$$

Notice from the definition of J in equation (17) 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 (C.2) and (C.3) 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.

C.1. Existence proof for the dynamic model

The main recursion, equation (B.25), 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 (B.25), 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 3. *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 $\|S\| \le \|T\|$ 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 $\{z \,\big|\, |z| \le \eta^{-1/2}\}$. 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 3 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

$$X_0 \equiv \left\{x : 0 \le |x| \le \left|x^*\right|\right\}.$$

This set is closed and bounded. The upper bound of $\left|T[X_0]\right|$ 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

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

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. □

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

To apply Lemma 3, 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:

$$T[D] \equiv \frac{1}{2} + \frac{D}{1+D}. \tag{C.4}$$

Also define the mapping associated in the recursion in (B.25) by S. We begin with:

Lemma 4.

$$|D| = |S[D]| = \left| \frac{1}{2} J^{-1} \left[J^{*-1} \sigma_e^2 \right]_+ + \frac{1}{2} J^{-1} \left[J \frac{2Re[D^*]}{1+D^*} \right]_+ \right| \le \frac{1}{2} + \left| \frac{D}{1+D} \right|. \tag{C.5}$$

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

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

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 5. *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

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

To establish boundedness, we must prove that $|1+D| > 1$ in the vicinity of the fixed point. We do this in Proposition 2.

Lemma 6. *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

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

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,

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

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

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 (B.31) and (B.32),

$$J^*J = 2\left[F^{*-1}D^*A\right]_+\left[F^{*-1}D^*A\right]_+^*(D^*+D+2D^*D)^{-1}\sigma_a^2+\sigma_e^2 \tag{C.6}$$

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 $A = \frac{1}{1-az}$):

$$\left[F^{*-1}D^*A\right]_+ = F(a)^{-1}D(a)A(z) = e^{-\frac{1}{2}\frac{1}{2\pi i}\oint\frac{\zeta+a}{\zeta-a}\ln\left(D^*D(D^*+D+2D^*D)^{-1}\right)\frac{d\zeta}{\zeta}}A(z) \tag{C.7}$$

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

To apply Lemma 3 we show that $D = 0$ is not a fixed point of the recursion S.

Lemma 7. $J(0) \neq \infty$.

Proof. Use the Szegö factorization to write

$$F^{-1}\left[F^{*-1}D^*A\right]_+ = e^{-\frac{1}{2}\frac{1}{2\pi i}\oint\frac{\zeta+a}{\zeta-a}\ln\left(D^*D(D^*+D+2D^*D)^{-2}\right)\frac{d\zeta}{\zeta}}A(z). \tag{C.8}$$

The inner term can be written as

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

$\frac{D}{D^*}$ 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. □

Lemma 8. *$D = 0$ 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 9. *There exists a lower bound $\underline{D}$ such that $|S[D]| > \underline{D}$ for all D.*

Proof. From Lemma 7, 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 $|\tilde{\theta}| < |\theta|$ 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 (C.4). The properties listed in Lemma 3 are satisfied for S and T by Lemmas 4, 5, 6.

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 (15) and the variational first-order condition (B.8), 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 (18) represents an optimum. The optimality of β and δ follow similarly. □

Proposition 2 in the main text follows.

Appendix D. Proofs of the characterisation results

Before proving Proposition 3 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 10. *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

$$\min_F E(A(L)e_t - F(L)(A(L)e_t + u_t))^2 = \min_F \frac{1}{2\pi i} \oint \left((A - FA)^*(A - FA)\sigma_e^2 + F^* F \sigma_u^2 \right) \frac{dz}{z}. \tag{D.1}$$

The variational first-order condition is

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

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

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 3. 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 (18) (the equation for α_i), as

$$\begin{aligned}\alpha_i &= F^{-1}A^{-1}\left[F^{*-1}(F^*F - D^*)A\right]_+ \\ &= F^{-1}A^{-1}\left([FA]_+ - \left[F^{*-1}D^*A\right]_+\right).\end{aligned} \tag{D.2}$$

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 (16) 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 (19), a scalar F and D means that J is driven by the A filter, which is exogenously non-scalar, a contradiction. □

Appendix E. State space methods in the numerical analysis

In order to numerically simulate and iterate the recursion in equation (20), 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

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

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,

$$y_t = Cx_t + Du_t \tag{E.2}$$

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 (E.1) 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 (E.2) 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 an element of the complex plane, which we denote z, and we can represent the process y_t simply by its z-transform,

$$C(I - Az)^{-1}B + D \tag{E.3}$$

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,

$$C(zI - A)^{-1}B + D \tag{E.4}$$

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 (E.4) 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

$$\begin{aligned}&(C_1(zI - A_1)^{-1}B_1 + D_1) + (C_2(zI - A_2)^{-1}B_2 + D_2)\\ &\qquad = \begin{pmatrix} C_1 & C_2 \end{pmatrix}\left(zI - \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}\right)^{-1}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} + \begin{pmatrix} D_1 + D_2 \end{pmatrix}\end{aligned} \tag{E.5}$$

which has the same basic form as (E.4). 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

$$\left[\begin{array}{cc|c} A_1 & B_1C_2 & B_1D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1C_2 & D_1D_2 \end{array}\right].$$

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

The details of these and other operation can be found in Dullerud and Paganini (2000), p. 99, or in Sanchez-Pena and Sznaier (1998), 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 (20).

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

E.1. The algorithm

I implemented these operations using Mathematica in order to numerically approximate a fixed point of equation (20), 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 (B.31) where a spectral factorization is carried out to calculate F, using the method devised in Taub (2009), 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 (B.32) 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 (20) 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 (20), 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 (2009). 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 C.

Appendix F. Proof of the inconspicuousness result

Proposition 9. *In equilibrium each firm sees the rival's net information in price as noise.*

Proof. The net information in price is, from equation (4),

$$\frac{1}{1+2\delta}\left((1-\alpha(L))A(L)a_{-it}+e_t\right) \tag{F.1}$$

Recalling the definition of D from (13) and J in (17), this expression becomes

$$DJ \tag{F.2}$$

in the frequency domain formulation. Using the recursive equation for D in (20), this becomes

$$\begin{aligned} DJ &= \left(\frac{1}{2}J^{-1}\left[J^{*-1}\sigma_e^2\right]_+ + \frac{1}{2}J^{-1}\left[J\frac{D+D^*}{1+D^*}\right]_+\right)J \\ &= \frac{1}{2}\left[J^{*-1}\sigma_e^2\right]_+ + \frac{1}{2}\left[J\frac{D+D^*}{1+D^*}\right]_+ \end{aligned} \tag{F.3}$$

The first term, $\frac{1}{2}\left[J^{*-1}\sigma_e^2\right]_+$, is clearly a scalar. The second term has only scalar terms because $D + D*$ is the real part in which all powers of z cancel each other, and the negative powers of z in $\frac{1}{1+D^*}$ are cancelled by the annihilator operator $[\cdot]_+$. □

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: High noise.

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: Public demand process persistent relative to private process.

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	1.0	**0.1**	**0.5**

Table 5
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
$0.02-\frac{2.10}{1.z-12.17}$	$0.008-\frac{1.08}{z-10.04}$

Table 6
Base Parametrization: Private demand process persistent relative to public process.

Discount factor η	Private variance σ_a^2	Public variance $\sigma_{\underline{a}}^2$	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 7
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 $\beta(z)$
$0.53+\frac{0.20}{z-2.19}-\frac{0.17}{z-22.}$	$0.31-\frac{0.007}{z-4.}+\frac{0.08}{z-2.19}-\frac{3.21}{z-110.}$
Total output on A	Total output on B
$0.05-\frac{0.77}{z-2.19}-\frac{0.31}{z-2.}-\frac{0.0003}{1.z-2.62}$	$0.06-\frac{0.035}{z-4.0}+\frac{0.0002}{z-3.11}-\frac{0.065}{z-2.19}-\frac{6.20}{z-10.0}$
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 8
Monte Carlo serial correlation: Example 3 parameterization.

Process	Output Autocorrelation	Price Autocorrelation
Direct private	0.41	
Indirect private	0.78	
Total private	0.52	0.50
Direct public	0.06	
Indirect public	0.55	
Total public	0.14	0.09
Noise process	0.36	0.17
Total if no noisy feedback from price	0.45	0.46
Total with noisy feedback	0.41	0.02
Net information in noisy price		-0.05

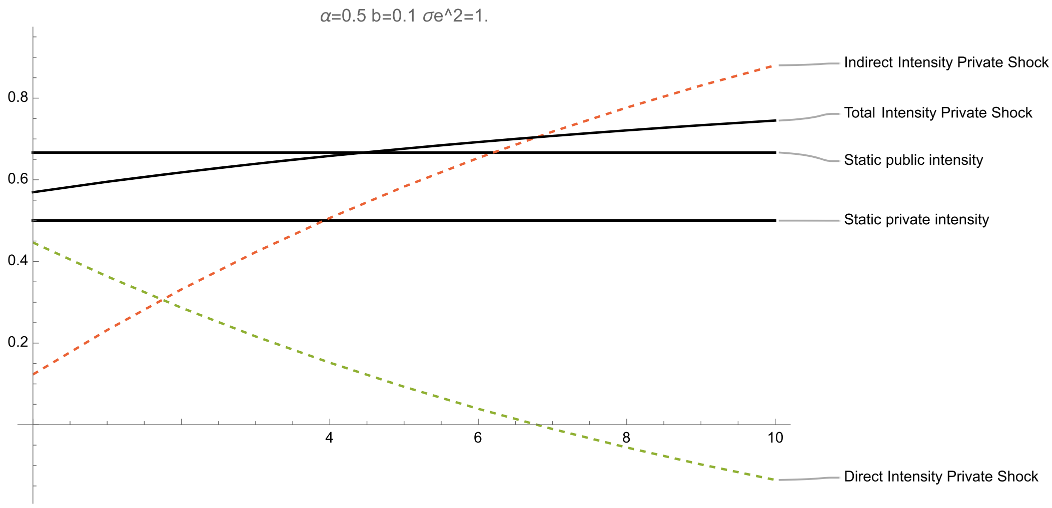

Fig. 1. Normalized intensities for private shock process $A(L)a_{it}$.

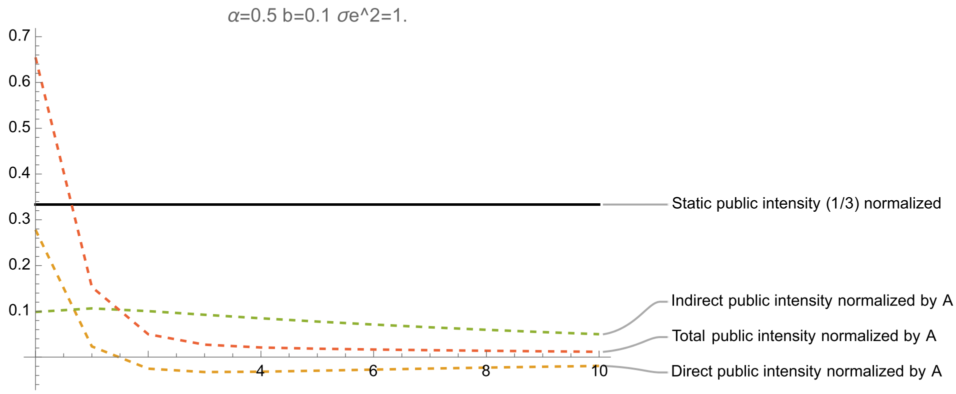

Fig. 2. Normalized intensities for public shock process $B(L)\overline{a}_{it}$.

References

Abreu, D., Pearce, D., Stacchetti, E., 1986. Optimal cartel equilibria with imperfect monitoring. J. Econ. Theory 39, 251–269.
Aghion, P., Bolton, C.H.P., Jullien, B., 1991. Optimal learning by experimentation. Rev. Econ. Stud. 58, 621–654.
Aghion, P., Espinosa, M., Jullien, B., 1993. Dynamic duopoly with learning through market experimentation. Econ. Theory 3 (3), 517–539.
Alepuz, D., Urbano, A., 2005. Learning in asymmetric duopoly markets: competition in information and market correlation. Span. Econ. Rev. 7 (3), 1435–5469.
Athey, S., Bagwell, K., 2008. Collusion with persistent cost shocks. Econometrica 76 (3), 493–540.
Bagwell, K., Ramey, G., 1991. Oligopoly limit pricing. Rand J. Econ. 22, 155–172.
Ball, J., Gohberg, I., Rodman, L., 1990. Interpolation of Rational Matrix Functions. Birkhäuser Verlag.
Bergemann, D., Heumann, T., Morris, S., 2015. Information and volatility. J. Econ. Theory 158, 427–465.
Bergin, J., Bernhardt, D., 2008. Industry dynamics with stochastic demand. Rand J. Econ., 41–68.
Bernhardt, D., Taub, B., 2015. Learning about common and private values in oligopoly. Rand J. Econ. 46 (1), 66–85.
Bernhardt, D., Seiler, P., Taub, B., 2010. Speculative dynamics. Rand J. Econ., 1–52.
Bonatti, A., Cisterna, G., Toikka, J., 2017. Dynamic oligopoly with incomplete information. Rev. Econ. Stud. 84 (2), 503–546.
Caminal, R., 2017. A dynamic duopoly model with asymmetric information. J. Ind. Econ. 38 (3), 315–333.
Caminal, R., Vives, X., 2017. Why market shares matter: an information-based theory. Rand J. Econ. 27 (2), 221–239.
Conway, J.B., 1985. A Course in Functional Analysis. Springer-Verlag.
Davenport, W.B., Root, W.L., 1958. An Introduction to the Theory of Random Signals and Noise. McGraw-Hill.
Dullerud, G., Paganini, F., 2000. A Course in Robust Control Theory. Springer.
Foster, D., Viswanathan, S., 1996. Strategic trading when agents forecast the forecasts of others. Journal of Finance LI (4), 1437–1478.
Green, E.J., Porter, R.H., 1984. Noncooperative collusion under imperfect price information. Econometrica 52 (1), 87–100.
Hackbarth, D., Taub, B., 2022. Does the potential to merge reduce competition. Manag. Sci. 68 (7), 5364–5383.
Hansen, L., Sargent, T., 1980. Formulating and estimating dynamic linear rational expectations models. J. Econ. Dyn. Control 2.
Harrington, J., 1986a. Limit pricing when the potential entrant is uncertain of its cost function. Econometrica 54 (2), 429–437.
Harrington, J., 1986b. Oligopolistic entry deterrence under incomplete information. Rand J. Econ. 18 (2), 211–231.
Harrington, J., 1995. Experimentation and learning in a differentiated-products duopoly. J. Econ. Theory 66 (1), 275–288.
Hoffman, K., 1962. Banach Spaces of Analytic Functions. Prentice-Hall.
Huo, Z., Takayama, N., 2023. Rational expectations models with higher order beliefs. SSRN. https://ssrn.com/abstract=3873663.
Kasa, K., 2000. Forecasting the forecasts of others in the frequency domain. Rev. Econ. Dyn. 3 (4), 726–756.

Kasa, K., Walker, T., Whiteman, C., 2014. Heterogeneous beliefs and tests of present value models. Rev. Econ. Stud. 81, 1137–1163.
Keller, G., Rady, S., 1999. Optimal experimentation in a changing environment. Rev. Econ. Stud. 66, 475–507.
Klemperer, P.D., Meyer, M.A., 1989. Supply function equilibria in oligopoly under uncertainty. Econometrica 57 (6), 1243–1277.
Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Kyle, A., 1989. Informed speculation with imperfect competition. Rev. Econ. Stud. 56, 317–355.
Mailath, G., 1989. Simultaneous signaling in an oligopoly model. Q. J. Econ. 104 (2), 417–427.
Makarov, I., Rytchkov, O., 1989. Forecasting the forecasts of others: implications for asset pricing. J. Econ. Theory, 941–966.
McLennan, A., 1984. Price dispersion and incomplete learning in the long run. J. Econ. Dyn. Control 7, 331–347.
Mirman, L.S., Urbano, A., 1993. Duopoly signal jamming. Econ. Theory 3 (1), 129–149.
Nimark, K., 2017. Dynamic higher order expectations. Mimeo.
Riordan, M., 1985. Imperfect information and dynamic conjectural variations. Rand J. Econ. 16 (1), 41–50.
Rondina, G., Walker, T.B., 2021. Confounding dynamics. J. Econ. Theory 196, 105251.
Rozanov, Y., 1967. Stationary Random Processes. Holden-Day.
Rudin, W., 1974. Real and Complex Analysis. McGraw-Hill.
Rustichini, A., Wolinsky, A., 1995. Learning about variable demand in the long run. J. Econ. Dyn. Control 19, 1283–1292.
Sanchez-Pena, R., Sznaier, M., 1998. Robust Systems: Theory and Applications. John Wiley.
Sannikov, Y., 2007. Games with imperfectly observable actions in continuous time. Econometrica 75 (5), 1285–1329.
Seiler, P., Taub, B., 2008. The dynamics of strategic information flows in stock markets. Finance Stoch. 12 (1), 43–82.
Taub, B., 1990. The equivalence of lending equilibria and signaling-based insurance under asymmetric information. Rand J. Econ. 21 (3), 388–408.
Taub, B., 2009. Implementing the Iakoubovski-Merino spectral factorization algorithm using state-space methods. Syst. Control Lett. 58 (6), 445–451.
Vives, X., 2011. Strategic supply competition with private information. Econometrica 79 (6), 1919–1966.
Whiteman, C., 1985. Spectral utility, Wiener-Hopf techniques, and rational expectations. J. Econ. Dyn. Control 9, 225–240.