2022

Essays On General Equilibrium

Sherwin Lott
University of Pennsylvania

Essays On General Equilibrium

Abstract

This dissertation develops new general equilibrium results on how markets react to risk and aggregate information.

The first chapter extends a classical result in portfolio theory about the effect of risk on value functions to its effect more generally on policy functions. If odd moments of shocks are zero up to some order, then the odd order marginal effects on value and policy functions of introducing these shocks are zero as well. Mathematically, all coefficients of corresponding odd order in the perturbation parameter are zero. If shocks are symmetric, e.g. normally distributed, then this holds for all odd orders. The main theorem (1) generalizes past results on perturbations and unifies their economic intuition, (2) improves the computation of stochastic coefficients, and (3) illustrates how to derive properties of high order perturbations through simple induction.

The second chapter tries to reconcile classical versions of the efficient market hypothesis with the surveyed level of technical analysis in practice. If past security prices are public information, then any patterns contained within should be approximately accounted for in current prices, and fundamental analysis would be relied on relatively more. While each past security price might individually be public information, the disconnect is that it should not imply their collective patterns and interactions are public information as well. Economics, unlike probability theory, must recognize costs and therefore distinguish between observing pieces of information and analyzing their many interactions. (1) We generalize sigma-fields to families of events and define information more broadly as knowledge about optimization solutions. (2) This provides a new framework for efficiency hypotheses and theorems. (3) We illustrate how complex patterns arise from "variably diffuse information" that only technical analysts can aggregate indirectly, changing the informational behavior of prices.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Economics

First Advisor

Jesus Fernandez-Villaverde

Subject Categories

Economics

ESSAYS ON GENERAL EQUILIBRIUM

Sherwin Lott

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2022

Supervisor of Dissertation

Jesús Fernández-Villaverde, Professor of Economics

Graduate Group Chairperson

David Dillenberger, Professor of Economics

Dissertation Committee

Andrew Shephard, Associate Professor of Economics
Andrew Postlewaite, Harry P. Kamen Professor of Economics

ACKNOWLEDGEMENT

I thank Jesús Fernández-Villaverde for introducing me to perturbation theory and posing the conjecture that motivated all of Chapter 1 in this dissertation. I greatly appreciate how caring and supportive Andrew Shephard has been throughout the program. Andrew Postlewaite has the rare depth and patience to take students' unformed ideas and hone in on the relevant economic intuition and big picture questions. I thank my committee for their time and latitude to follow my passion in Chapter 2.

ABSTRACT

ESSAYS ON GENERAL EQUILIBRIUM

Sherwin Lott

Jesús Fernández-Villaverde

This dissertation develops new general equilibrium results on how markets react to risk and aggregate information.

The first chapter extends a classical result in portfolio theory about the effect of risk on value functions to its effect more generally on policy functions. If odd moments of shocks are zero up to some order, then the odd order marginal effects on value and policy functions of introducing these shocks are zero as well. Mathematically, all coefficients of corresponding odd order in the perturbation parameter are zero. If shocks are symmetric, e.g. normally distributed, then this holds for all odd orders. The main theorem (1) generalizes past results on perturbations and unifies their economic intuition, (2) improves the computation of stochastic coefficients, and (3) illustrates how to derive properties of high order perturbations through simple induction.

The second chapter tries to reconcile classical versions of the efficient market hypothesis with the surveyed level of technical analysis in practice. If past security prices are public information, then any patterns contained within should be approximately accounted for in current prices, and fundamental analysis would be relied on relatively more. While each past security price might individually be public information, the disconnect is that it should not imply their collective patterns and interactions are public information as well. Economics, unlike probability theory, must recognize costs and therefore distinguish between observing pieces of information and analyzing their many interactions. (1) We generalize sigma-fields to families of events and define information more broadly as knowledge about optimization solutions. (2) This provides a new framework for efficiency hypotheses and theorems. (3) We

illustrate how complex patterns arise from “variably diffuse information" that only technical analysts can aggregate indirectly, changing the informational behavior of prices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

Perturbations in DSGE Models: An Odd Derivatives Theorem

1.1. Introduction

This first chapter is a reproduction of Lott (2019) with permission from Elseiver.

Dynamic stochastic general equilibrium (DSGE) models are a primary workhorse of macroeconomics. Since DSGE models are generally intractable, a variety of approaches have been developed to compute their solutions. One popular way is with *perturbation methods*, which are quite fast and yet maintain reasonable accuracy (Judd and Guu (1993); Judd (1998); Judd and Guu (2001); Jin and Judd (2002); Aruoba et al. (2006); Caldara et al. (2012)). This method approximates the DSGE solution by a multivariate Taylor series with respect to the state variables $\mathbf{x}$ and a *perturbation parameter* σ.

A growing number of applications require high order perturbations to capture the relevant effects. Fernandez-Villaverde et al. (2011) show that only perturbations of the third–order or higher can accurately model volatility shocks. de Groot (2015) and de Groot (2016) argue that perturbations as high as the fourth or sixth–order are necessary for asset pricing models with stochastic volatility. While Schmitt-Grohe and Uribe (2004) get tractable expressions for second–order perturbations, analytic results for higher orders are nearly impossible as they involve many high-dimensional tensor products. The development of new theoretical approaches and results for high order perturbations is increasingly essential.

Computation gets much more difficult with each successive order as well: runtime becomes a binding constraint, and estimation errors grow large. Swanson et al. (2006) demonstrate how tiny numerical errors rapidly compound on one another as higher order perturbations build on lower orders. They find that "coefficient errors as large as 10% become quite common by about the fourth or fifth order." Thus, there is a greater focus on the theory and computation of high order perturbations. For example, Levintal (2017) analytically develops a "new notation" that treats the perturbation parameter like a state variable. This approach allows a manifold increase in computational speed when implemented for fifth–order perturbations.

The direct contribution of our paper is to generalize a result in Jin and Judd (2002) to all orders. They prove that Taylor coefficients of the first order in the perturbation parameter σ are zero when shocks have mean zero.[1] (i.e. The coefficients of order $\mathbf{x}^r\sigma^1$ are zero for all r.) More recently, Andreasen (2012) proved that the $\mathbf{x}^0\sigma^3$ coefficient is zero if the first and third moment of the shocks are zero. Fernandez-Villaverde et al. (2016) goes on to conjecture and verify with examples that all coefficients of odd σ–order are zero if shocks are symmetric. We prove in Theorem 1 that: if all odd shock moments up to order $\bar{s}$ are zero, then coefficients of order $\mathbf{x}^r\sigma^s$ are zero for all r and odd $s \leq \bar{s}$. The prior results and conjecture are all special cases of our theorem. In fact, we get a considerably stronger version of Andreasen (2012) where all the cross derivatives with respect to the state variables are zero as well.

Theorem 1 is simply the natural extension to policy functions of a classical result in portfolio theory. A special case of this result is that investors with no risk in their portfolio are marginally risk neutral. Put in more abstract terms, the first order effect of risk on value functions is zero. The general result is that: the order s effect of risk on value functions is proportional to moment s of the underlying shock. If particular moments are zero, then the corresponding order effects of risk on value functions are zero as well. All our theorem does is extend these results on value functions to policy functions. If there is no order s effect of risk on value functions, then there must be no order s effect of risk on the value maximizing decisions. Since this holds over all state variables $\mathbf{x}$, any cross derivatives with respect to $\mathbf{x}$ must remain zero.

This extension of portfolio theory has direct implications for the computation of high order perturbations. For instance, all odd σ–order coefficients can be set precisely to zero if shocks

[1]Without loss of generality, shocks always have mean zero.

are symmetric (as they often are, e.g. normally distributed). The odd σ–order equations that determine these coefficients no longer need to be computed, and the even σ–order equations can be simplified. Thus, high order perturbations can be computed faster and with fewer numerical errors. These gains are permanent in the sense that they are not implementation dependent: zeroing out coefficients and symbolic derivatives is useful no matter what future techniques or notations are developed.

Beyond improving economic intuition and computation, the proof itself serves as a template for working with high order perturbations. Past theory has been constrained by the algebraic difficulty of tensor products that grow rapidly in complexity with each order. That is why proving just the third order result in Andreasen (2012) was a difficult feat. We cut through all of this by establishing a simple inductive pattern on the parity of derivative orders, which proves the equations have a particular form that implies odd σ–order coefficients are zero. Such patterns can be found by carefully inspecting the differentiation that determines the equations, and they give rise to simple proofs of properties that hold for all orders.

The remainder of this paper is organized as follows. Section 1.2 details the standard perturbation setup in macroeconomics and proves a fundamental property about its equations. Then, Section 3 applies this property inductively to prove the main theorem and generalize past results. Section 1.4 develops the underlying economic intuition through a portfolio theory example. Finally, Section 1.5 analyzes various computational improvements, and Section 2.6 concludes.

1.2. Perturbation Setup

The idea of a perturbation is to approximate solutions over the space of models, which then gives the solution to any particular model. This is just a Taylor series, but one in which the domain is over model parameters. We start with a simple model, in that there is a known or readily computable solution, and build an approximation around it.

In macroeconomics, we want to compute the equilibrium policy functions to a DSGE model.

Consider the space over all transformations of this DSGE model where shocks are scaled by a constant perturbation parameter $\sigma \geq 0$. The deterministic model corresponds to $\sigma = 0$, which has a relatively easy to compute steady–state. Policy functions are then approximated by a Taylor series with respect to the state variables $\mathbf{x}$ and the perturbation parameter σ that are centered at the steady–state. The equilibrium to our unscaled model, $\sigma = 1$, can then be backed out.

Section 1.2.1 details the standard perturbation setup within macroeconomics. Then, Section 1.2.2 derives a simple observation (Lemma 1) that is at the heart of the main proof in Section 1.3.

1.2.1. Standard Setup

We consider perturbations in the context of a generic DSGE model.[2] Denote the control variables of this model by $\mathbf{y}_t \in \mathbb{R}^{n_y}$ and the state variables by $\mathbf{x}_t \in \mathbb{R}^{n_x}$. Let the equilibrium conditions be expressible as a system of equations:

$$\mathbb{E}_t \mathcal{H} \left(\mathbf{y}_t, \mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{x}_{t+1} \right) = 0 \tag{1.1}$$

Denote the corresponding policy functions by $\mathbf{g}$ and $\mathbf{h}$, where:

$$\mathbf{y}_t = \mathbf{g} \left(\mathbf{x}_t; \sigma \right), \tag{1.2}$$

$$\mathbf{x}_{t+1} = \mathbf{h} \left(\mathbf{x}_t; \sigma \right) + \sigma \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+1} \tag{1.3}$$

The state variables evolve stochastically. Here, $\boldsymbol{\epsilon}_{t+1}$ is a vector of n_ϵ independent shocks; $\boldsymbol{\eta}$ is a $n_x \times n_\epsilon$ matrix—it linearly transforms the shocks into state variable shocks.[3] Note that this setup allows for any state variable covariance matrix through the choices of $\boldsymbol{\eta}$ and $\boldsymbol{\epsilon}_{t+1}$. Finally, the perturbation parameter $\sigma \geq 0$ is a constant that scales the magnitude of

[2] We follow the generalized setup in Fernandez-Villaverde et al. (2016).

[3] While η is a constant at time t, it may depend on past $\boldsymbol{\epsilon}$'s in any way. Thus allowing for stochastic volatility and wide variety of other dependencies.

the shocks.

When these shocks are scaled away, $\sigma = 0$, the variables evolve deterministically. Denote the deterministic steady-state of the model as $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, satisfying:

$$\mathcal{H}\left(\bar{\mathbf{y}}, \bar{\mathbf{y}}, \bar{\mathbf{x}}, \bar{\mathbf{x}}\right) = 0 \tag{1.4}$$

Solving this yields the fixed point of the policy functions: $\bar{\mathbf{y}} = \mathbf{g}\left(\bar{\mathbf{x}}, 0\right)$ and $\bar{\mathbf{x}} = \mathbf{h}\left(\bar{\mathbf{x}}, 0\right)$.

We want to back out the policy functions to the unscaled model, $\sigma = 1$, which are $\mathbf{g}\left(\,\cdot\,, 1\right)$ and $\mathbf{h}\left(\,\cdot\,, 1\right)$. These are approximated by computing Taylor series for $\mathbf{g}$ and $\mathbf{h}$ centered at the deterministic steady-state.

The coefficients of this Taylor series are computed through an implicit function argument. Using decision rules (1.2) and (1.3), the variables $\mathbf{y}_t$, $\mathbf{y}_{t+1}$, and $\mathbf{x}_{t+1}$ can be expressed in terms of $\mathbf{x}_t$ and σ. Plugging these into the equilibrium condition gives an expression that only depends on $\mathbf{x}_t$ and σ, denote this as the function F:

$$F(\mathbf{x};\sigma) \equiv \mathbb{E}\left[\mathcal{H}\Big(\mathbf{g}(\mathbf{x};\sigma),\, \mathbf{g}\big(\mathbf{h}(\mathbf{x};\sigma) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon};\sigma\big),\, \mathbf{x},\, \mathbf{h}(\mathbf{x};\sigma) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}\Big)\right] = 0$$

By construction, F is always zero; hence, any derivatives of F must also evaluate to zero, $F_{\mathbf{x}^r\sigma^s}(\mathbf{x};\sigma) = 0$. Taking these derivatives up to some finite n^{th}–order, forms a system of equations where the unknowns are derivatives of $\mathbf{g}$ and $\mathbf{h}$ to the n^{th}–order. This system has exactly as many equations as unknowns, which allows the Taylor expansion to be computed.

1.2.2. Derivatives of F

The derivatives of F form a system of equations that determine the policy functions $\mathbf{g}$ and $\mathbf{h}$. Understanding the functional form of these derivatives is key. As an example, consider the first σ–order derivative:[4]

$$F_{\mathbf{x}^0\sigma^1}(\mathbf{x};\sigma) = \mathbb{E}\left[\mathcal{H}_{\mathbf{y}}\mathbf{g}_\sigma + \mathcal{H}_{\mathbf{y}'}\hat{\mathbf{g}}_{\mathbf{x}}\mathbf{h}_\sigma + \mathcal{H}_{\mathbf{y}'}\hat{\mathbf{g}}_{\mathbf{x}}\boldsymbol{\eta}\boldsymbol{\epsilon} + \mathcal{H}_{\mathbf{y}'}\hat{\mathbf{g}}_\sigma + \mathcal{H}_{\mathbf{x}'}\mathbf{h}_\sigma + \mathcal{H}_{\mathbf{x}'}\boldsymbol{\eta}\boldsymbol{\epsilon}\right]$$

Any order derivative is readily obtained by repeatedly applying the product and chain rules. Multiplying these expressions out, they can be represented generically as a sum of products.

Claim 1. *The equations $F_{\mathbf{x}^r\sigma^s}$ can be expressed as a finite sum of products:*

$$F_{\mathbf{x}^r\sigma^s}(\mathbf{x};\sigma) = \mathbb{E}\left[\sum_{j=1}^{J^{(r,s)}} P^{(r,s,j)}\right]$$

Each $P^{(r,s,j)}$ is the product of terms of the form: $\mathcal{H}_\bullet$, $\mathbf{g}_\bullet$, $\hat{\mathbf{g}}_\bullet$, $\mathbf{h}_\bullet$, and $\boldsymbol{\eta}\boldsymbol{\epsilon}$.[5] The total number of products being summed over is $J^{(r,s)}$, with j indexing over these products.

One approach would be to write out each $P^{(r,s,j)}$ explicitly, but this would give unnecessarily complex tensor products. Instead, for our purposes, we only need to show that when s is odd, $P^{(r,s,j)}$ contains an odd number of $\boldsymbol{\epsilon}$'s or an odd σ–order derivative of $\mathbf{g}$ or $\mathbf{h}$. To do this, we will show that the sum over the frequency of $\boldsymbol{\epsilon}$ and σ–orders of derivatives of $\mathbf{g}$ and $\mathbf{h}$ is s. (This summation will be made more precise.) When s is odd, there must then be at least one such odd term.

The idea is that when $P^{(r,s,j)}$ is differentiated by σ, each new product will contain exactly either one more $\boldsymbol{\epsilon}$ or another σ–order derivative of $\mathbf{g}$ or $\mathbf{h}$. When $P^{(r,s,j)}$ is differentiated by $\mathbf{x}$, there will be no new such terms. This will inductively show that the previously mentioned sum is s, which will be sufficient information about the products to prove the theorem.

To prove this idea, we need to determine how each $P^{(r,s,j)}$ changes when differentiated. When $P^{(r,s,j)}$ is differentiated, the product rule will split it up into a sum of products where

[4] With the appropriate shorthand notations: $\mathbf{g}_\bullet = \mathbf{g}_\bullet(\mathbf{x};\sigma)$, $\hat{\mathbf{g}}_\bullet = \mathbf{g}_\bullet\big(\mathbf{h}(\mathbf{x};\sigma) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon};\sigma\big)$, $\mathbf{h}_\bullet = \mathbf{h}_\bullet(\mathbf{x};\sigma)$, and $\mathcal{H}_\bullet = \mathcal{H}_\bullet\Big(\mathbf{g}(\mathbf{x};\sigma),\, \mathbf{g}\big(\mathbf{h}(\mathbf{x};\sigma) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon};\sigma\big),\, \mathbf{x},\, \mathbf{h}(\mathbf{x};\sigma) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}\Big)$. Dropping time subscripts, $\mathbf{x}'$ and $\mathbf{y}'$ are shorthand for next period variables.

[5] This holds inductively. The derivative of any one of these terms is an expression containing only these terms multiplied or added together.

only one term is being differentiated. Exhaustively, the following are all the derivatives that arise from differentiating one term in a product:

(With respect to σ.)

(D1) $\frac{\partial}{\partial\sigma}\mathcal{H}_\bullet = \mathcal{H}_{\bullet\mathbf{y}}\mathbf{g}_\sigma + \mathcal{H}_{\bullet\mathbf{y}'}\hat{\mathbf{g}}_\mathbf{x}\mathbf{h}_\sigma + \mathcal{H}_{\bullet\mathbf{y}'}\hat{\mathbf{g}}_\mathbf{x}\boldsymbol{\eta\epsilon} + \mathcal{H}_{\bullet\mathbf{y}'}\hat{\mathbf{g}}_\sigma + \mathcal{H}_{\bullet\mathbf{x}'}\mathbf{h}_\sigma + \mathcal{H}_{\bullet\mathbf{x}'}\boldsymbol{\eta\epsilon}$

(D2) $\frac{\partial}{\partial\sigma}\mathbf{g}_{\mathbf{x}^r\sigma^s} = \mathbf{g}_{\mathbf{x}^r\sigma^{s+1}}$

(D3) $\frac{\partial}{\partial\sigma}\hat{\mathbf{g}}_{\mathbf{x}^r\sigma^s} = \hat{\mathbf{g}}_{\mathbf{x}^{r+1}\sigma^s}\mathbf{h}_\sigma + \hat{\mathbf{g}}_{\mathbf{x}^{r+1}\sigma^s}\boldsymbol{\eta\epsilon} + \hat{\mathbf{g}}_{\mathbf{x}^r\sigma^{s+1}}$

(D4) $\frac{\partial}{\partial\sigma}\mathbf{h}_{\mathbf{x}^r\sigma^s} = \mathbf{h}_{\mathbf{x}^r\sigma^{s+1}}$

(With respect to $\mathbf{x}$.)

(D5) $\frac{\partial}{\partial\mathbf{x}}\mathcal{H}_\bullet = \mathcal{H}_{\bullet\mathbf{y}}\mathbf{g}_\mathbf{x} + \mathcal{H}_{\bullet\mathbf{y}'}\hat{\mathbf{g}}_\mathbf{x}\mathbf{h}_\mathbf{x} + \mathcal{H}_{\bullet\mathbf{x}} + \mathcal{H}_{\bullet\mathbf{x}'}\mathbf{h}_\mathbf{x}$

(D6) $\frac{\partial}{\partial\mathbf{x}}\mathbf{g}_{\mathbf{x}^r\sigma^s} = \mathbf{g}_{\mathbf{x}^{r+1}\sigma^s}$

(D7) $\frac{\partial}{\partial\mathbf{x}}\hat{\mathbf{g}}_{\mathbf{x}^r\sigma^s} = \hat{\mathbf{g}}_{\mathbf{x}^{r+1}\sigma^s}\mathbf{h}_\mathbf{x}$

(D8) $\frac{\partial}{\partial\mathbf{x}}\mathbf{h}_{\mathbf{x}^r\sigma^s} = \mathbf{h}_{\mathbf{x}^{r+1}\sigma^s}$

Notice that the changes in these terms are in line with what we had expected. When a term is differentiated by σ, (D1)–(D4), each new product has exactly either an additional $\boldsymbol{\epsilon}$ or another σ–order derivative of $\mathbf{g}$ or $\mathbf{h}$. Whereas, when a term is differentiated by $\mathbf{x}$, (D5)-(D8), each new product has no such additional $\boldsymbol{\epsilon}$'s or σ–order derivatives.

To be rigorous, we explicitly track these term totals. For each $P^{(r,s,j)}$, we denote the number of $\boldsymbol{\epsilon}$ terms and define a sum over the σ–orders of derivatives of $\mathbf{g}$ and $\mathbf{h}$. These are expressed in the following notation as $a^{(r,s,j)}$ and $b^{(r,s,j)}$ respectively.

Notation 1.

- *Let* $a^{(r,s,j)}$ *denote the total number of* $\boldsymbol{\epsilon}$*'s that are multiplied in* $P^{(r,s,j)}$.
- *Let* $k_i^{(r,s,j)}$ *denote the* σ*–order of the derivative for the* i^{th} $\mathbf{g}$ *or* $\mathbf{h}$ *term in* $P^{(r,s,j)}$*, and*

denote the sum by, $b^{(r,s,j)} = \sum_i k_i^{(r,s,j)}$.

We have argued that these should sum to s, as stated in the following lemma. This will be the core of the proof—the only piece of information needed about the products to prove our theorem.

Lemma 1. $a^{(r,s,j)} + b^{(r,s,j)} = s, \;\; \forall r, s, j.$

Proof. This holds for the base case when $r = 0$ and $s = 0$.

Inductively, suppose this holds for some r, s: $a^{(r,s,j)} + b^{(r,s,j)} = s$, $\forall j$. Then, consider each new term obtained from the product rule on $\frac{\partial}{\partial \sigma} P^{(r,s,j)}$. Inspecting (D1) - (D4), any of these differentiations increases $a+b$ by one in any new products. Hence, $a^{(r,s+1,j)}+b^{(r,s+1,j)} = s+1$, $\forall j$.

Similarly, consider each new term obtained from the product rule on $\frac{\partial}{\partial \mathbf{x}} P^{(r,s,j)}$. Inspecting (D5) - (D8), none of these differentiations effect a or b. Hence, $a^{(r+1,s,j)} + b^{(r+1,s,j)} = s$. This lemma then holds by induction. □

This implies that if s is odd, then either $a^{(r,s,j)}$ or $b^{(r,s,j)}$ is odd. Further, then $P^{(r,s,j)}$ must have an odd σ–order derivative or an odd number of $\boldsymbol{\epsilon}$'s.

1.3. Main Theorem

This paper partial solves the odd σ–order coefficients in the following theorem.[6]

Theorem 1. *If* $\mathbb{E}[\boldsymbol{\epsilon}^s] = 0$ *for all odd* $s \leq \bar{s}$*, then* $\mathbf{g}_{\mathbf{x}^r \sigma^s}(\bar{\mathbf{x}}; 0) = \mathbf{h}_{\mathbf{x}^r \sigma^s}(\bar{\mathbf{x}}; 0) = 0$*, for all* r *and odd* $s \leq \bar{s}$*.*

The proof is laid out in Section 1.3.1 and followed by a discussion in Section 1.3.2 of how this theorem generalizes past results in the literature. Then, Section 1.4 and Appendix 1.7 go through the intuition behind it using an example from classical portfolio theory.

[6] Let $\boldsymbol{\epsilon}^s$ denote the Kronecker product of $\boldsymbol{\epsilon}$ with itself s times.

1.3.1. Proof of Theorem

This paper is about proving a specific property of perturbations, so we take as given that the standard methods in Section 1.2 are mathematically valid. That means the system of equations is well defined—derivatives of $\mathcal{H}$ exist to whatever order is taken. And, in particular, we assume there exists a unique solution to each perturbation order. Without these assumptions, which are implicitly made in previous papers, perturbation methods should not be used in the first place. Conditions under which these assumptions hold are outside the purview of this paper. (See Lan and Meyer-Gohde (2014) for solvability conditions.)

Keeping the parity pattern of Lemma 1 in mind, we now turn to partially solving the system of equations generated by $F_{\mathbf{x}^r\sigma^s} = 0$.[7] First, it should be emphasized, the role that evaluating $F_{\mathbf{x}^r\sigma^s}$ at the deterministic steady-state plays. Having $\sigma = 0$ eliminates all of the $\boldsymbol{\epsilon}$ terms within functions, so all the functions can be treated as constants with respect to expectations. In conjunction, evaluating at $\bar{\mathbf{x}}$ makes $\mathbf{g}$ and $\hat{\mathbf{g}}$ equivalent—all the functions are being evaluated at the deterministic steady-state.

The idea behind our partial solution is that, in a n^{th}–order perturbation, odd σ–order equations are solved by setting odd σ–order unknowns $\mathbf{g}_{\mathbf{x}^r\sigma^s}$ and $\mathbf{h}_{\mathbf{x}^r\sigma^s}$ to zero. This will follow from Lemma 1—every product in the odd σ–order equations contains either a zero or an odd number of $\boldsymbol{\epsilon}$'s.

Lemma 2. *If* $\mathbb{E}[\boldsymbol{\epsilon}^s] = 0$ *for all odd* $s \leq \bar{s}$, *and* $\mathbf{g}_{\mathbf{x}^r\sigma^s}(\bar{\mathbf{x}}; 0) = \mathbf{h}_{\mathbf{x}^r\sigma^s}(\bar{\mathbf{x}}; 0) = 0$ *for all* r *and odd* $s \leq \bar{s}$ *where* $r + s \leq n$*; then, the equations* $F_{\mathbf{x}^r\sigma^s}(\bar{\mathbf{x}}; 0) = 0$ *are satisfied for all such previous* r *and* s.

Proof. Take any such r and s. By Lemma 1, every product in $F_{\mathbf{x}^r\sigma^s}$ multiplies an odd σ-order ($\leq \bar{s}$) policy derivative or an odd number ($\leq \bar{s}$) of $\boldsymbol{\epsilon}$ terms.[8] All other terms in

[7] All functions are evaluated at the deterministic steady state $(\bar{\boldsymbol{x}}; 0)$ in this section.

[8] The ordering of the $\boldsymbol{\epsilon}$'s does not matter. Every product in the resultant tensor will contain that many elements of $\boldsymbol{\epsilon}$, which evaluates to zero in expectation if there is an odd such number $\leq \bar{s}$.

$F_{\mathbf{x}^r\sigma^s}$ are constant because functions are being evaluated at the deterministic steady-state. Hence, each product in $F_{\mathbf{x}^r\sigma^s}$ evaluates to zero in expectation. □

The number of equations eliminated is equal to the number of unknowns being set. Theorem 1 can now be proven by induction.

Proof of Theorem 1. This is proven by induction on the perturbation order.
Schmitt-Grohe and Uribe (2004) already proved the base case that $\mathbf{g}_\sigma = \mathbf{h}_\sigma = 0$.

Suppose this theorem holds for a n^{th}–order perturbation, we want to show it then holds for the next $(n+1)^{th}$–order perturbation. By Lemma 2, setting the $(n+1)^{th}$–order policy derivatives corresponding with this theorem to zero eliminates exactly as many equations as unknowns. We now invoke the fact that the $(n+1)^{th}$–order equations are linear given the solution to the n^{th}–order perturbation. This concludes the proof. Our proposed partial solution eliminates as many equations as unknowns in a linear system with a unique solution; therefore, it is in fact a partial solution. □

1.3.2. Generalizing Past Results

Theorem 1 generalizes the previous results in the literature: Jin and Judd (2002), Schmitt-Grohe and Uribe (2004), Andreasen (2012), and proves an open conjecture in Fernandez-Villaverde et al. (2016). We illustrate this by writing them out as corollaries for when $\bar{s} = 1$, 3, and ∞.

Corollary 1. *If* $\mathbb{E}[\boldsymbol{\epsilon}] = 0$*, then* $\mathbf{g}_{\mathbf{x}^r\sigma} = \mathbf{h}_{\mathbf{x}^r\sigma} = 0$*, for all* r*.*[9]

Corollary 1 was first proven in Theorem 7 of the manuscript Jin and Judd (2002). Schmitt-Grohe and Uribe (2004) concurrently proved a weaker result while deriving tractable solutions to second order perturbations.

[9] The assumption of mean zero shocks is without loss of generality.

Corollary 2. *If* $\mathbb{E}[\boldsymbol{\epsilon}^3] = 0$*, then* $\mathbf{g}_{x^r\sigma^3} = \mathbf{h}_{x^r\sigma^3} = 0$*, for all* r*.*

Andreasen (2012) proved that $\mathbb{E}[\boldsymbol{\epsilon}^3] = 0$ implies $\mathbf{g}_{\sigma^3} = \mathbf{h}_{\sigma^3} = 0$, which Corollary 2 extends to all possible cross derivatives with respect to the state variables.

Corollary 3. *If* $\boldsymbol{\epsilon}$ *is symmetric, then* $\mathbf{g}_{\mathbf{x}^r\sigma^s} = \mathbf{h}_{\mathbf{x}^r\sigma^s} = 0$*, for all* r *and odd* s*.*

Corollary 3 was conjectured and verified by Fernandez-Villaverde et al. (2016), page 559. This corollary is widely applicable because symmetric shocks are common (e.g. normal distributions). We detail in Section 1.5 how these results can be used to improve computational methods.

1.4. Intuition from Portfolio Theory

Theorem 1 is actually just the natural extension of a classical result in portfolio theory,[10] which thus summarizes all the intuition behind our theorem. We illustrate this by breaking down a portfolio example into two parts: (1) the effect of risk on value functions and (2) how that translates into the effect of risk on policy functions. The effect on value functions is already well known, and our contribution is in extending that result to policy functions. The example here is continued to higher orders in Appendix 1.7.

1.4.1. Risk and Value Functions

Consider a two-period model with an investor who has wealth w and earns a rate of return $r + \sigma\epsilon$.[11] (w, r, and σ are constants, and ϵ is a random variable.) The investor's expected utility from choosing an investment level y is given by:

$$v(y \,|\, \sigma) \equiv u(c_1) + \mathbb{E}\left[u(c_2)\right]$$

where $c_1 = w - y$ and $c_2 = y(1 + r + \sigma\epsilon)$.

[10] The earliest reference we can find for this result is Pratt (1964).

[11] The example can be rewritten using log returns to avoid negative consumption. This would not qualitatively change the results that follow.

Derivatives of the value function v with respect to risk σ are readily computed:

$$\frac{\partial v}{\partial \sigma}(y \,|\, \sigma) = \mathbb{E}\left[u(c_2)y\epsilon\right] \qquad \Longrightarrow \frac{\partial v}{\partial \sigma}(y \,|\, 0) = u(c_2)y\mathbb{E}\left[\epsilon\right]$$

$$\frac{\partial^s v}{\partial \sigma^s}(y \,|\, \sigma) = \mathbb{E}\left[u(c_2)y^s\epsilon^s\right] \qquad \Longrightarrow \frac{\partial^s v}{\partial \sigma^s}(y \,|\, 0) = u(c_2)y^s\mathbb{E}\left[\epsilon^s\right]$$

Second period consumption c_2 is deterministic when evaluated at $\sigma = 0$, which leads the order s derivative to simply scale with moment s.

We reinterpret this result as the following:

Remark 1. *The order s effect of risk on value functions is proportional to moment s (when moving from a deterministic state).*

For example, investors are marginally risk neutral when they have no risk in their portfolio. (In other words, there is no first order effect of risk since $\mathbb{E}\left[\epsilon\right] = 0$.) This is the intuition given by Jin and Judd (2002) for Corollary 1.

As an implication of remark 1, if moment s is zero, then derivative s is also zero.

Remark 2. *For any s:*

$$\mathbb{E}\left[\epsilon^s\right] = 0 \quad \Longrightarrow \quad \frac{\partial^s v}{\partial \sigma^s}(y \,|\, 0) = 0 \quad \forall y$$

Since this derivative is zero for all y, any cross derivatives are zero as well.

$$\mathbb{E}\left[\epsilon^s\right] = 0 \quad \Longrightarrow \quad \frac{\partial^{s+1} v}{\partial y \partial \sigma^s}(y \,|\, 0) = 0 \quad \forall y$$

While Remark 2 holds for all orders s, even and odd moments are fundamentally different. Even moments are always positive since they measure how spread out the shock is, whereas odd moments are often zero and measure asymmetries in the shock.[12] Thus, we focus

[12]Economists generally use symmetric shocks and rather do appropriate transformations of it within the model itself. Indicative of this is that the macroeconomic software Dynare only allows perturbations with

exclusively on how the odd order terms can be zeroed out.

1.4.2. Risk and Policy Functions

Let $y^*(\sigma)$ denote the optimal level of investment, which is characterized by the first order condition (FOC):[13]

$$\frac{\partial v}{\partial y}\left(y^*(\sigma)\,|\,\sigma\right) = 0$$

A natural followup question is whether Remark 2 extends to policy functions. This would be rather intuitive since policy functions simply characterize the utility maximizing decisions. Thus, if there is no order s effect of risk on value functions, then there should be no order s effect of risk on policy functions.

To prove this intuition, we must compute the derivatives of $y^*(\sigma)$ by applying implicit function arguments. The first derivative is given by the equation:

$$\frac{\partial}{\partial\sigma}\frac{\partial v}{\partial y}\left(y^*(\sigma)\,|\,\sigma\right) = 0 \quad \Longrightarrow \quad \left[\frac{\partial^2 v}{\partial y^2}\frac{\partial y^*}{\partial\sigma} + \frac{\partial^2 v}{\partial y\partial\sigma}\right]_{(y^*(\sigma),\,\sigma)} = 0$$

If $\mathbb{E}[\epsilon] = 0$, then $\frac{\partial^2 v}{\partial y \partial\sigma}$ is zero when $\sigma = 0$, and therefore $\frac{\partial y^*}{\partial\sigma}(0) = 0$.

Higher order equations are computed in Appendix 1.7.[14] For example, the third order equation can be expressed as follows:

$$\left[(\ldots)\frac{\partial^3 y^*}{\partial\sigma^3} + (\ldots)\frac{\partial y^*}{\partial\sigma} + (\ldots)\frac{\partial^4 v}{\partial y\partial\sigma^3} + (\ldots)\frac{\partial^2 v}{\partial y\partial\sigma}\right]_{(y^*(\sigma),\,\sigma)} = 0$$

If $\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon^3] = 0$, then $\frac{\partial^2 v}{\partial y\partial\sigma}$, $\frac{\partial y^*}{\partial\sigma}$, and $\frac{\partial^4 v}{\partial y\partial\sigma^3}$ are zero when $\sigma = 0$. This in turn implies by the third order equation that $\frac{\partial^3 y^*}{\partial\sigma^3}(0) = 0$.

A pattern arises for all orders about how the odd order equations can be expressed. It is

normally distributed shocks.

[13] Assuming as we do for perturbation methods that there exists a unique solution in the interior. The FOC is analogous to the equilibrium conditions F.

[14] While this is a straightforward application of the product and chain rules, the expressions become long and tedious.

that every product contains a function being differentiated by an odd σ–order. We prove this using an elementary parity argument about how odd integers cannot be expressed as sums of even integers. (Even integers can, however, be expressed as sums of odd integers.)

Appealing to this pattern, the following remark is proven by induction.

Remark 3. *If* $\mathbb{E}[\epsilon^s] = 0$ *for all odd* $s \leq \bar{s}$*, then* $\frac{\partial^s y^*}{\partial \sigma^s}(0) = 0$ *for all odd* $s \leq \bar{s}$*.*

An example of the inductive step is when we used $\mathbb{E}[\epsilon] = 0 \Rightarrow \frac{\partial y^*}{\partial \sigma}(0) = 0$ to conclude that $\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon^3] = 0 \Rightarrow \frac{\partial^3 y^*}{\partial \sigma^3}(0) = 0$.

1.4.3. Special Case of the Main Theorem

Remarks 1 and 2 are classical results from portfolio theory about the effect of risk on value functions, and Remark 3 is their natural extension to policy functions. While this extension is relatively difficult to prove (as it involves induction via an implicit function argument), its intuition is the same as the portfolio theory result that was trivial to derive.

The proof of Theorem 1 is analogous to what was done in this portfolio example. That is, Remarks 1 and 2 hold in general because the perturbation parameter σ linearly scales the shocks $\boldsymbol{\epsilon}$ by construction. The first half of the proof simply observes that equations have the particular parity pattern between σ–order derivatives and shock moments (Lemma 1). And finally, Lemma 2 proves the inductive step.

This portfolio example is, in fact, just a rather simple perturbation. Appendix 1.7 rewrites the example as such and analytically computes the system of equations corresponding to the fourth order perturbation.

1.5. Computational Improvements

The computational complexity and memory requirements for perturbations become onerous as the number of coefficients grows *exponentially* in perturbation order and state variables (Appendix 1.9.1). Further, numerical errors compound as high order perturbations builds on lower order solutions (Swanson et al. (2006)). This section develops various direct ways that our results help improve computation.

Theorem 1 proved that odd σ–order coefficients are zero when corresponding odd shock moments are zero. (If $\boldsymbol{\epsilon}$ is symmetric, then all odd σ–order coefficients are zero.) Lemma 2 proved this partial solution solves the corresponding odd σ–order equations, which allows their symbolic differentiation to be avoided altogether (Section 1.5.1). Thus, a significant fraction of coefficients and equations can be eliminated as in Table 1, which produces an even greater reduction in runtime (Appendix 1.8).

Perturbation Order

Number of State Variables n_x \ n	1	2	3	4	5	10	20
1	.500	.400	.444	.429	.450	.462	.478
2	.333	.333	.368	.382	.400	.439	.466
3	.250	.286	.324	.348	.368	.420	.456
4	.200	.250	.291	.320	.343	.403	.446
5	.167	.222	.265	.297	.321	.388	.436
10	.091	.143	.185	.218	.245	.326	.394
20	.048	.083	.115	.142	.165	.247	.330
50	.020	.037	.054	.069	.083	.142	.221

Table 1: Proportion of coefficients that are of odd σ–order.
(See Appendix 1.9.3 for more details.)

Moreover, the even σ–order coefficients and equations that remain are easier to compute (Appendix 1.7.3). These equations are sums of long products, and any product that contains an eliminated coefficient is itself constantly zero over the state variables.[15] Setting these products to zero greatly simplifies the equations (Appendix 1.7.3), and should similarly improve computation.[16]

1.5.1. Modified Block Recursion

This section explains in detail how the *block recursion* method from Jin and Judd (2002) can be improved. Let *block–s* refer to the set of equations of order σ^s, which in Figure 1 is depicted by the entire row s. The blocks are computed recursively, starting with block–0.

[15] If product $P^{(r,s,j)}$ contains a coefficient of odd σ–order less than or equal to $\bar{s}$, then $P^{(r,s,j)} = 0$.

[16] We have implemented code that removes odd σ–order coefficients and equations (Appendix 1.8). Though, we have not implemented this improvement of the even equations.

Algorithm 1 (Modified Block Recursion)**.** *Suppose* $\mathbb{E}[\boldsymbol{\epsilon}^s] = 0$ *for all odd* $s \leq \bar{s}$*, then:*

1. *Compute the deterministic block-0 normally.*
2. *Compute* $F_{\mathbf{x}^0\sigma^2}$*, then set odd* σ*–order coefficients and odd moments* ($\leq \bar{s}$) *to zero (thus eliminating any products they are in). Denote this new expression by* $\hat{F}_{\mathbf{x}^0\sigma^2}$*.*
3. *Compute block-2 using* $\hat{F}_{\mathbf{x}^0\sigma^2}$*.*
4. *Repeat steps 2 and 3 for each even block–s (and for any odd block-s where* $s > \bar{s}$*).*

Claim 2. *By construction,* $\frac{\partial^r}{\partial \mathbf{x}^r} F_{\mathbf{x}^0\sigma^s}(\mathbf{x}; 0) = \frac{\partial^r}{\partial \mathbf{x}^r} \hat{F}_{\mathbf{x}^0\sigma^s}(\mathbf{x}; 0)$*.*

Proof. All the products removed from $F_{\mathbf{x}^0\sigma^s}(\mathbf{x}; 0)$ are constantly zero over $\mathbf{x}$. The odd σ–order coefficients ($\leq \bar{s}$) are zero for all $\mathbf{x}$ by Theorem 1, and the shock moments are unaffected by differentiating with respect to $\mathbf{x}$ by (D5)–(D8). □

Therefore, block-s can be computed using $\hat{F}_{\mathbf{x}^0\sigma^s}$. See Appendix 1.7.3 for an analytic example of $\hat{F}$ and how it substantially reduces symbolic differentiation.

$$
\begin{array}{llllllll}
⓪ & F_{\mathbf{x}^0\sigma^0} \longrightarrow & F_{\mathbf{x}^1\sigma^0} \longrightarrow & F_{\mathbf{x}^2\sigma^0} \longrightarrow & \cdots \longrightarrow & F_{\mathbf{x}^{n-2}\sigma^0} \longrightarrow & F_{\mathbf{x}^{n-1}\sigma^0} \longrightarrow & F_{\mathbf{x}^n\sigma^0} \\
& \downarrow & & & & & & \\
① & F_{\mathbf{x}^0\sigma^1} & F_{\mathbf{x}^1\sigma^1} & F_{\mathbf{x}^2\sigma^1} & \cdots & F_{\mathbf{x}^{n-2}\sigma^1} & F_{\mathbf{x}^{n-1}\sigma^1} & \\
& \downarrow & & & & & & \\
② & F_{\mathbf{x}^0\sigma^2} \longrightarrow & \hat{F}_{\mathbf{x}^1\sigma^2} \longrightarrow & \hat{F}_{\mathbf{x}^2\sigma^2} \longrightarrow & \cdots \longrightarrow & \hat{F}_{\mathbf{x}^{n-2}\sigma^2} & & \\
& \downarrow & & & & & & \\
③ & F_{\mathbf{x}^0\sigma^3} & F_{\mathbf{x}^1\sigma^3} & F_{\mathbf{x}^2\sigma^3} & \cdots & & & \\
& \downarrow & & & & & & \\
& \vdots & & \cdots & & & & \\
& \downarrow & & & & & & \\
ⓝ & F_{\mathbf{x}^0\sigma^n} & & & & & &
\end{array}
$$

Figure 1: Modified block recursion.

1.6. Conclusion

This paper extends a classical result about the effect of risk on value functions to their effect more generally on policy functions. If the shock moments of odd order $s \leq \bar{s}$ are zero, then there are no odd order $s \leq \bar{s}$ effects of risk on value or policy functions at the deterministic steady-state. The theorem generalizes prior results in the perturbation literature and unifies their economic intuition around a basic concept in portfolio theory. This also permanently improves computation by reducing coefficients and equations. When shocks are symmetric (as they commonly are), all odd σ–order coefficients and equations are eliminated, and even σ–order equations are substantially simplified.

Beyond these particular results, our inductive proof provides a novel template for deriving properties of perturbations that apply to all orders. Rather than manipulating equations for each particular order, we instead focus on finding patterns within the differentiation process itself. Future research could, for example, explain patterns in the now simplified even σ–order equations (Appendix 1.7.3). Theoretical progress has up until now been severely constrained by how algebraically complicated tensor products become with each additional order. Since applications are becoming more sophisticated and require high order perturbations, research in this area is increasingly essential.

1.7. Appendix: Portfolio Theory Example

This Appendix continues the example from Section 1.4 and computes all the implicit function equations up to the fourth order. Investors have a value function $v(y \,|\, \mathbf{x}, \sigma)$ for investment y, and we take as given that Remark 2 holds:

$$\mathbb{E}\left[\epsilon^s\right] = 0 \quad \Longrightarrow \quad \frac{\partial^s v}{\partial \sigma^s}(y \,|\, \mathbf{x}, 0) = 0 \;\; \forall y \text{ and } \mathbf{x}$$

The optimal level of investment is denoted by $y^*(\mathbf{x}, \sigma)$, and satisfies the first order condition:

$$F(\mathbf{x}; \sigma) \equiv \frac{\partial v}{\partial y}\left(y^*(\mathbf{x}, \sigma) \,|\, \mathbf{x}, \sigma\right) = 0$$

1.7.1. Derivatives

We can now proceed to take derivatives of F with respect to σ. The first order equation ($F_{\mathbf{x}^0\sigma^1} = 0$) computed in Section 1.4 is:

$$\left[\frac{\partial^2 v}{\partial y^2}\frac{\partial y^*}{\partial \sigma} + \frac{\partial^2 v}{\partial y \partial \sigma}\right]_{(y^*(\mathbf{x},\sigma),\,\mathbf{x},\,\sigma)} = 0$$

Here we have two "products," and each has *one* derivative with respect to σ.

Differentiation

While taking derivatives is a straightforward application of the product and chain rules, the expressions do get long and tedious. There are only two functions here, v and y^*, so the reader should keep in mind that their derivatives are:

$$\frac{\partial}{\partial \sigma}\left[\frac{\partial^{r+s} v}{\partial y^r \sigma^s}\right]_{(y^*(\mathbf{x},\sigma),\,\mathbf{x},\,\sigma)} = \left[\frac{\partial^{r+s+1} v}{\partial y^{r+1}\partial \sigma^s}\frac{\partial y^*}{\partial \sigma} + \frac{\partial^{r+s+1} v}{\partial y^r \sigma^{s+1}}\right]_{(y^*(\mathbf{x},\sigma),\,\mathbf{x},\,\sigma)} \tag{1.1}$$

and,

$$\frac{\partial}{\partial \sigma}\left[\frac{\partial^{s} y^*}{\partial \sigma^s}\right]_{(\mathbf{x},\,\sigma)} = \left[\frac{\partial^{s+1} y^*}{\partial \sigma^{s+1}}\right]_{(\mathbf{x},\,\sigma)} \tag{1.2}$$

Our simple observation is that: when taking the derivative with respect to σ using the product rule, each resulting product will have one more derivative with respect to σ. For notational brevity we will drop the evaluation subscript $(y^*(\mathbf{x},\sigma),\,\mathbf{x},\,\sigma)$.

Second Order Derivative

The second derivative of F with respect to σ yields the equation $F_{\mathbf{x}^0\sigma^2} = 0$:

$$\begin{aligned}\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial y^*}{\partial \sigma}\right)^2 + 2\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial y^*}{\partial \sigma} + \frac{\partial^3 v}{\partial y \partial \sigma^2} & \\ + \frac{\partial^2 v}{\partial y^2}\frac{\partial^2 y^*}{\partial \sigma^2} & = 0\end{aligned}$$

There are four separate products and each has *two* derivatives with respect to σ. Two of the four products contain a first order derivative in σ (either $\frac{\partial y^*}{\partial \sigma}$ or $\frac{\partial^3 v}{\partial y^2 \partial \sigma}$).

Third Order Derivative

The third derivative of F with respect to σ yields the equation $F_{\mathbf{x}^0\sigma^3} = 0$:

$$\begin{aligned}\frac{\partial^4 v}{\partial y^4}\left(\frac{\partial y^*}{\partial \sigma}\right)^3 + 3\frac{\partial^4 v}{\partial y^3 \partial \sigma}\left(\frac{\partial y^*}{\partial \sigma}\right)^2 + 3\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial y^*}{\partial \sigma} + \frac{\partial^4 v}{\partial y \partial \sigma^3} \\ + 3\frac{\partial^3 v}{\partial y^3}\frac{\partial^2 y^*}{\partial \sigma^2}\frac{\partial y^*}{\partial \sigma} + 3\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial^2 y^*}{\partial \sigma^2} \\ + \frac{\partial^2 v}{\partial y^2}\frac{\partial^3 y^*}{\partial \sigma^3} && = 0\end{aligned}$$

There are seven separate products and each has *three* derivatives with respect to σ. Every product contains an odd derivative in σ (either $\frac{\partial y^*}{\partial \sigma}$, $\frac{\partial^3 y^*}{\partial \sigma^3}$, $\frac{\partial^3 v}{\partial y^2 \partial \sigma}$, $\frac{\partial^4 v}{\partial y \partial \sigma^3}$, or $\frac{\partial^4 v}{\partial y^3 \partial \sigma}$).

Fourth Order Derivative

The fourth derivative of F with respect to σ yields the equation ($F_{\mathbf{x}^0\sigma^4} = 0$):

$$\begin{aligned}\frac{\partial^5 v}{\partial y^5}\left(\frac{\partial y^*}{\partial \sigma}\right)^4 + 4\frac{\partial^5 v}{\partial y^4 \partial \sigma}\left(\frac{\partial y^*}{\partial \sigma}\right)^3 + 6\frac{\partial^5 v}{\partial y^3 \partial \sigma^2}\left(\frac{\partial y^*}{\partial \sigma}\right)^2 + 4\frac{\partial^5 v}{\partial y^2 \partial \sigma^3}\frac{\partial y^*}{\partial \sigma} + \frac{\partial^5 v}{\partial y \partial \sigma^4} \\ + 6\frac{\partial^4 v}{\partial y^4}\frac{\partial^2 y^*}{\partial \sigma^2}\left(\frac{\partial y^*}{\partial \sigma}\right)^2 + 12\frac{\partial^4 v}{\partial y^3 \partial \sigma}\frac{\partial^2 y^*}{\partial \sigma^2}\frac{\partial y^*}{\partial \sigma} + 6\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial^2 y^*}{\partial \sigma^2} \\ + 4\frac{\partial^3 v}{\partial y^3}\frac{\partial^3 y^*}{\partial \sigma^3}\frac{\partial y^*}{\partial \sigma} + 3\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial^2 y^*}{\partial \sigma^2}\right)^2 + 4\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial^3 y^*}{\partial \sigma^3} \\ + \frac{\partial^2 v}{\partial y^2}\frac{\partial^4 y^*}{\partial \sigma^4} && = 0\end{aligned}$$

There are twelve separate products and each has *four* derivatives with respect to σ. Eight of the twelve products contain a first order derivative in σ (either $\frac{\partial y^*}{\partial \sigma}$, $\frac{\partial^3 y^*}{\partial \sigma^3}$, $\frac{\partial^3 v}{\partial y^2 \partial \sigma}$, $\frac{\partial^4 v}{\partial y^3 \partial \sigma}$, $\frac{\partial^5 v}{\partial y^2 \partial \sigma^3}$, or $\frac{\partial^5 v}{\partial y^4 \partial \sigma}$).

1.7.2. Odd Order Solutions

With the previous equations computed, the derivatives of y^* can now be solved for.

First Order Solution

Suppose that $\mathbb{E}[\epsilon] = 0$, then $\frac{\partial v}{\partial \sigma}(y \mid \mathbf{x}, 0)$ is constantly zero over y and $\mathbf{x}$ (by Remark 2). The first order equation $F_{\mathbf{x}^0\sigma^1}(\mathbf{x}, 0) = 0$ then becomes:[17]

$$\begin{aligned} &\left[\frac{\partial^2 v}{\partial y^2}\frac{\partial y^*}{\partial \sigma} + \xcancel{\frac{\partial^2 v}{\partial y \partial \sigma}}\right]_{(y^*(\mathbf{x},0),\, \mathbf{x},\, 0)} = 0 \\ &\implies \frac{\partial y^*}{\partial \sigma}(\mathbf{x}, 0) = 0 \end{aligned}$$

As this holds for all $\mathbf{x}$, it must be that:

$$\frac{\partial^{r+1} y^*}{\partial \mathbf{x}^r \partial \sigma}(\mathbf{x}, 0) = 0 \quad \forall r$$

Alternatively, differentiating $F_{\mathbf{x}^0\sigma^1}$ with respect to $\mathbf{x}$ does not affect the σ–order of any derivatives. Thus, $\frac{\partial^{r+1} y^*}{\partial \mathbf{x}^r \partial \sigma}(\mathbf{x}, 0) = 0$ for all r solves the equations $F_{\mathbf{x}^r\sigma^1} = 0$ for all r.

Third Order Solution

Similarly, suppose that $\mathbb{E}[\epsilon] = 0$ and $\mathbb{E}\left[\epsilon^3\right] = 0$, and therefore: $\frac{\partial v}{\partial \sigma}(y \mid \mathbf{x}, 0)$, $\frac{\partial y^*}{\partial \sigma}(\mathbf{x}, 0)$, and $\frac{\partial^3 v}{\partial \sigma^3}(y \mid \mathbf{x}, 0)$ are constantly zero over all y and $\mathbf{x}$. The third order equation $F_{\mathbf{x}^0\sigma^3}(\mathbf{x}, 0) = 0$ then becomes:

$$\begin{aligned} &\xcancel{\frac{\partial^4 v}{\partial y^4}\left(\frac{\partial y^*}{\partial \sigma}\right)^3} + 3\xcancel{\frac{\partial^4 v}{\partial y^3 \partial \sigma}\left(\frac{\partial y^*}{\partial \sigma}\right)^2} + 3\xcancel{\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial y^*}{\partial \sigma}} + \xcancel{\frac{\partial^4 v}{\partial y \partial \sigma^3}} \\ &+ 3\xcancel{\frac{\partial^3 v}{\partial y^3}\frac{\partial^2 y^*}{\partial \sigma^2}\frac{\partial y^*}{\partial \sigma}} + 3\xcancel{\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial^2 y^*}{\partial \sigma^2}} \\ &+ \frac{\partial^2 v}{\partial y^2}\frac{\partial^3 y^*}{\partial \sigma^3} = 0 \end{aligned}$$

$$\implies \frac{\partial^3 y^*}{\partial \sigma^3}(\mathbf{x}, 0) = 0$$

[17]We assume the solution is unique.

Again, since this holds for all $\mathbf{x}$, it must be that:

$$\frac{\partial^{r+3} y^*}{\partial \mathbf{x}^r \partial \sigma^3}(\mathbf{x}, 0) = 0 \quad \forall r$$

This pattern repeats itself to all odd σ–orders because the number of σ–derivatives within each product is odd. Therefore, each product must have at least one odd σ–order derivative of v or y^*.

1.7.3. Even Order Solutions

The even order equations are substantially simplified by the fact that odd σ–order terms are zero. This simplifying can be done before taking any cross derivatives with respect to the state variables $\mathbf{x}$, and thus reduces symbolic differentiations.

Second Order Solution

With the odd σ–order coefficients set to zero, the derivative $\hat{F}_{\mathbf{x}^0 \sigma^2}$ becomes:[18]

$$\begin{aligned} &\xcancel{\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial y^*}{\partial \sigma}\right)^2} + 2\xcancel{\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial y^*}{\partial \sigma}} + \frac{\partial^3 v}{\partial y \partial \sigma^2} \\ &\quad + \frac{\partial^2 v}{\partial y^2}\frac{\partial^2 y^*}{\partial \sigma^2} \\ &= \frac{\partial^3 v}{\partial y \partial \sigma^2} + \frac{\partial^2 v}{\partial y^2}\frac{\partial^2 y^*}{\partial \sigma^2} \end{aligned}$$

The equation $\hat{F}_{\mathbf{x}^0 \sigma^2}(\mathbf{x}; 0) = 0$ implies that:

$$\frac{\partial^2 y^*}{\partial \sigma^2} = -\frac{\partial^3 v}{\partial y \partial \sigma^2} \Bigg/ \frac{\partial^2 v}{\partial y^2}$$

[18] We are using the $\hat{F}$ notation defined in Algorithm 1.

Fourth Order Solution

Similarly, $\hat{F}_{\mathbf{x}^0\sigma^4}$ becomes:

$$
\begin{aligned}
&\xcancel{\frac{\partial^5 v}{\partial y^5}\left(\frac{\partial y^*}{\partial \sigma}\right)^4} + 4\xcancel{\frac{\partial^5 v}{\partial y^4 \partial \sigma}\left(\frac{\partial y^*}{\partial \sigma}\right)^3} + 6\xcancel{\frac{\partial^5 v}{\partial y^3 \partial \sigma^2}\left(\frac{\partial y^*}{\partial \sigma}\right)^2} + 4\xcancel{\frac{\partial^5 v}{\partial y^2 \partial \sigma^3}\frac{\partial y^*}{\partial \sigma}} + \frac{\partial^5 v}{\partial y \partial \sigma^4} \\
&\quad + 6\xcancel{\frac{\partial^4 v}{\partial y^4}\frac{\partial^2 y^*}{\partial \sigma^2}\left(\frac{\partial y^*}{\partial \sigma}\right)^2} + 12\xcancel{\frac{\partial^4 v}{\partial y^3 \partial \sigma}\frac{\partial^2 y^*}{\partial \sigma^2}\frac{\partial y^*}{\partial \sigma}} + 6\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial^2 y^*}{\partial \sigma^2} \\
&\quad + 4\xcancel{\frac{\partial^3 v}{\partial y^3}\frac{\partial^3 y^*}{\partial \sigma^3}\frac{\partial y^*}{\partial \sigma}} + 3\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial^2 y^*}{\partial \sigma^2}\right)^2 + 4\xcancel{\frac{\partial^3 v}{\partial y^2 \partial \sigma}\frac{\partial^3 y^*}{\partial \sigma^3}} \\
&\quad + \frac{\partial^2 v}{\partial y^2}\frac{\partial^4 y^*}{\partial \sigma^4} \\
&= \frac{\partial^5 v}{\partial y \partial \sigma^4} + 6\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial^2 y^*}{\partial \sigma^2} + 3\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial^2 y^*}{\partial \sigma^2}\right)^2 + \frac{\partial^2 v}{\partial y^2}\frac{\partial^4 y^*}{\partial \sigma^4}
\end{aligned}
$$

The equation $\hat{F}_{\mathbf{x}^0\sigma^4}(\mathbf{x};0) = 0$ implies that:

$$
\frac{\partial^4 y^*}{\partial \sigma^4} = -\left[\frac{\partial^5 v}{\partial y \partial \sigma^4} + 6\frac{\partial^4 v}{\partial y^2 \partial \sigma^2}\frac{\partial^2 y^*}{\partial \sigma^2} + 3\frac{\partial^3 v}{\partial y^3}\left(\frac{\partial^2 y^*}{\partial \sigma^2}\right)^2\right] \Big/ \frac{\partial^2 v}{\partial y^2}
$$

It is intriguing how $\frac{\partial^2 v}{\partial y^2}$ scales the even σ–order coefficients.

1.8. Appendix: Simple Coding Example

This Appendix does a partial[19] implementation of Algorithm 1. The PerturbationAIM[20] algorithm in Swanson et al. (2006) is modified so that odd σ–order coefficients and equations are no longer computed, though even σ–order equations are left unchanged.[21] Computational improvements are then proportional to the coefficients eliminated in Table 1.

1.8.1. Coding Details

All of the code herein is originally from Swanson et al. (2006), Rudebusch and Swanson (2008), and Levintal (2017). I chose to modify the PerturbationAIM version 2.8.3 because its Mathematica code is clear and concise. The only changes made are to the “AIMStDeriva-

[19]This Appendix is meant as a proof of concept. The full modification of established perturbation solvers is left to more experienced programmers.

[20]This code is maintained by Eric Swanson and can be found on his personal webpage.

[21]I thank Eric Swanson for showing me how to modify the code. I take full responsibility for its correctness as the final changes are my own.

tives" function.[22] Three different versions are compared in the following section:

1. All equations are computed the "standard" way.

2. The results in Jin and Judd (2002) are used to skip the first σ–order equations. ("Skip first," which is what PerturbationAIM v2.8.3 currently does.)

3. Our results are used to skip odd σ–order equations as in Figure 1. ("Skip odd.")

1.8.2. Levintal (2017)—Two State Variables

This section measures the runtime reductions from eliminating stochastic[23] coefficients and equations as previously detailed. Tables 2 and 3 analyze these runtimes in the models of Levintal (2017) (two state variables) and Rudebusch and Swanson (2008) (six state variables).[24]

Table 2 gives the runtime (in seconds) of computing perturbations up to the fifth order.[25] This is broken down into the deterministic part (unchanged) and the stochastic part (computed in three different ways).

		2	3	4	5
Deterministic		.40	3.69	33.68	321.50
Stochastic	Standard	.12	3.61	47.71	692.18
	Skip First	.05	.50	12.83	256.13
	Skip Odd	.05	.41	7.31	156.46

Table 2: Runtimes of PerturbationAIM.

For example, it took a total of $321.50 + 692.18 = 1013.68$ seconds to compute a fifth order perturbation the standard way.

Table 3 gives the improvements as a proportion of the standard runtime.

For example, the proportional improvement in runtime for the fifth order perturbation of

[22] One other line of code is changed: "stdrvindxs" gives the new number of stochastic coefficients.

[23] The *deterministic* coefficients and equations are those of zeroth σ–order, whereas *stochastic* refers to those of nonzero σ–order. No changes are made to the deterministic computation.

[24] We use these models because their respective codes "rssmodel.m" and "rbc_with_disasters.m" were formatted for PerturbationAIM. Shocks in Levintal (2017) are changed to be symmetric.

[25] We used Mathematica 11.2.0.0 with macOS Mojave 10.14 on a 24GB iMac.

	2	3	4	5
Skip First	.138	.427	.428	.430
Skip Odd	.134	.439	.496	.528

Table 3: Reduction in runtime (proportion).

skipping all the odd σ–order coefficients and equations (using Table 2) is: $1 - (321.50 + 156.46)/1013.68 = .528$.

Lastly, Table 4 gives the computational improvement of skipping all odd σ–orders relative to just skipping the first σ–order.

3	4	5
.023	.119	.173

Table 4: Relative reduction in runtime.

For example, the relative reduction in runtime for a fifth order perturbation (using Table 3) is: $1 - (1 - .528)/(1 - .430) = .173$.

1.8.3. Rudebusch and Swanson (2008)—Six State Variables

All the tables in Section 1.8.2 are replicated here for Rudebusch and Swanson (2008) up to the fourth order. Runtimes increase dramatically when there are many more state variables (six versus two).

		2	3	4
Deterministic		1.18	34.07	1413.72
Stochastic	Standard	1.72	57.14	2745.97
	Skip First	.29	19.30	1235.83
	Skip Odd	.36	18.31	1180.88

Table 5: Runtimes of PerturbationAIM.

1.8.4. Conclusion

Tables 3 and 6 give considerably higher reductions in runtime than the proportion of coefficients eliminated in Table 1. This is because σ affects all state variables that have shocks. Therefore, derivatives with respect to σ are more complicated than derivatives with respect

	2	3	4
Skip First	.493	.415	.363
Skip Odd	.469	.426	.376

Table 6: Reduction in runtime (proportion).

3	4
.018	.021

Table 7: Relative reduction in runtime.

to any individual state variable (see (D1)–(D8)). Counting coefficients in the following Appendix 1.9 then gives an underestimate.

More state variables make the lower σ–order parts of a perturbation relatively difficult to compute. That is why Table 4 shows relatively large computational gains compared with Table 7. However, even with skipping all odd σ–orders, the stochastic portion of the perturbation is computationally intensive. Thus, there is a lot of room to improve the stochastic even σ–orders (For example, we have yet to implement code that zeroes out products in the even σ–order equations using Claim 2. This would substantially reduce symbolic differentiations as shown in Appendix 1.7.3.)

1.9. Appendix: Counting Coefficients

Let us quantify how many coefficients there are and what proportion are of odd σ–order. This measures the computational complexity of perturbations and the degree to which our results can simplify them.

1.9.1. Total Number of Coefficients in a Perturbation

Consider a perturbation of order n with n_y control variables and n_x state variables. The following Claim 3 gives an exact closed form expression for the number of coefficients in such an order n perturbation.

Claim 3. *The number of coefficients to be estimated in an order n perturbation is:*

$$T(n, n_x, n_y) = (n_x + n_y)\left(\binom{n + n_x + 1}{n_x + 1} - 1\right)$$

Proof. The policy functions $\mathbf{g}$ and $\mathbf{h}$ are of length n_y and n_x respectively. The total number of coefficients is the number of ways these policy functions can be differentiated (in an order n perturbation) multiplied by $n_x + n_y$.

All derivatives of these policy functions must be computed up to order n with respect to the state variables $\mathbf{x} = (x_1, \ldots x_{n_x})$ and the perturbation parameter σ. The number of such derivatives is the number of expressions $\partial x_1^{r_1} \partial x_2^{r_2} \ldots \partial x_{n_x}^{r_{n_x}} \partial \sigma^s$ with nonnegative integer powers satisfying $1 \leq r_1 + r_2 + \ldots + r_{n_x} + s \leq n$. By a standard "stars and bars" combinatorial argument, this is $\binom{n+n_x+1}{n_x+1} - 1$. □

Using this closed form expression, we can say exactly how the total number of coefficients is affected by n, n_x, and n_y. Using properties of binomials, we can say that the total number of coefficients:

- Is linear in the number of control variables n_y (fixing n and n_x).
- Behaves (asymptotically) like the polynomial n_x^{n+1} (fixing n and n_y).
- Behaves (asymptotically) like the polynomial n^{n_x+1} (fixing n_x and n_y).
- Grows exponentially as n and n_x are multiplied by a scaling c (fixing n_y).

1.9.2. Proportion of Coefficients that are of Odd σ–Order

We want to compute the proportion of coefficients that are of odd σ–order. For this, it is more illustrative to think of $T(n, n_x, n_y)$ as the following sum.

$$T(n, n_x, n_y) = (n_x + n_y)\left(\sum_{0 \leq s \leq n} \binom{n - s + n_x}{n_x} - 1\right)$$

Here, we are summing over the number of coefficients with σ–order s. For a fixed s, the number of nonnegative integer solutions to $r_1 + r_2 + \ldots + r_{n_x} \leq n - s$ is, $\binom{(n-s)+n_x}{n_x}$. (Again, by a "stars and bars" combinatorial argument.)

Now, we can express the number of odd σ–order coefficients.

$$T_{odd}(n, n_x, n_y) = (n_x + n_y) \sum_{\substack{\text{odd } s, \\ 0 \leq s \leq n}} \binom{n - s + n_x}{n_x}$$

Hence, the proportion of coefficients of odd σ–order is:

$$p_{odd}(n, n_x, n_y) = T_{odd}(n, n_x, n_y)/T(n, n_x, n_y)$$

To simplify notation, we drop n_y.

$$p_{odd}(n, n_x) = T_{odd}(n, n_x)/T(n, n_x)$$

Where:

$$T_{odd}(n, n_x) = T_{odd}(n, n_x, n_y)/(n_x + n_y)$$

$$T(n, n_x) = T(n, n_x, n_y)/(n_x + n_y)$$

Claim 4. *The proportion of coefficients of odd σ–order is bounded between:*

$$p_{odd}(n, n_x) \;\leq\; \frac{n}{2n + n_x} \underbrace{\left(1 + \frac{1}{T(n, n_x)}\right)}_{\approx 1 \text{ (large } n \text{ or } n_x)}$$

$$p_{odd}(n, n_x) \;\geq\; \underbrace{\left(\frac{n}{2(n-1) + n_x}\right)}_{\approx \frac{n}{2n+n_x} \text{ (large } n \text{ or } n_x)} \underbrace{\left(\frac{n + n_x - 1}{n + n_x + 1}\right)\left(1 + \frac{1}{T(n, n_x)}\right)}_{\approx 1 \text{ (large } n \text{ or } n_x)}$$

This claim should be interpreted as $p_{odd}(n, n_x) \approx \frac{n}{2n+n_x}$.

Proof. First, we'll simplify p_{odd} by eliminating the -1 in its denominator. (This is simply coming from $T(n, n_x)$ not counting the intercepts of the policy functions.) Denote,

$$\hat{p}_{odd}(n, n_x) = p_{odd}(n, n_x)\left(1 + \frac{1}{T(n, n_x)}\right)^{-1}$$

Now,

$$\hat{p}_{odd}(n, n_x) = \sum_{\substack{\text{odd } s, \\ 0 \leq s \leq n}} \binom{n - s + n_x}{n_x} \Big/ \sum_{0 \leq s \leq n} \binom{n - s + n_x}{n_x}$$

The critical part of the argument is that the terms in the summation are *decaying faster than geometrically.*

$$\binom{n - (s+1) + n_x}{n_x} \Big/ \binom{n - s + n_x}{n_x} = \frac{n - s}{n - s + n_x}$$

The first term in the summation corresponds with $s = 0$, which is even. Each even term is followed by an odd term of proportion less than $\frac{n}{n+n_x}$. This then gives a lower bound on the sum of the even terms.

$$\frac{n}{n + n_x} \sum_{\substack{\text{even } s, \\ 0 \leq s \leq n}} \binom{n - s + n_x}{n_x} \geq \sum_{\substack{\text{odd } s, \\ 0 \leq s \leq n}} \binom{n - s + n_x}{n_x}$$

This then gives an upper bound on $\hat{p}_{odd}(n, n_x)$:

$$\begin{aligned} \hat{p}_{odd}(n, n_x) &\leq \left(1 + \frac{n}{n + n_x}\right)^{-1} \\ &= \frac{n}{2n + n_x} \end{aligned}$$

The lower bound is proven in the same way after accounting for the first term ($s = 0$). The first term is the following percentage of the sum.

$$\binom{n-0+n_x}{n_x} \Big/ \sum_{0\le s\le n} \binom{n-s+n_x}{n_x} = \binom{n+n_x}{n_x} \Big/ \binom{n+n_x+1}{n_x+1}$$
$$= \frac{n_x+1}{n+n_x+1}$$

This implies that the sum over $1 \le s \le n$ is,

$$\sum_{1\le s\le n} \binom{n-s+n_x}{n_x} = \left(1 - \frac{n_x+1}{n+n_x+1}\right) \sum_{0\le s\le n} \binom{n-s+n_x}{n_x}$$
$$= \underbrace{\frac{n}{n+n_x+1}}_{} \sum_{0\le s\le n} \binom{n-s+n_x}{n_x}$$

Now, each odd term is followed by an even term of proportion less than $\frac{n-1}{n-1+n_x}$. This then gives the lower bound:

$$\frac{n-1}{n-1+n_x} \sum_{\substack{\text{odd } s,\\ 1\le s\le n}} \binom{n-s+n_x}{n_x} \ge \sum_{\substack{\text{even } s,\\ 1\le s\le n}} \binom{n-s+n_x}{n_x}$$

$$\implies \sum_{\substack{\text{odd } s,\\ 1\le s\le n}} \binom{n-s+n_x}{n_x} \ge \left(1 + \frac{n-1}{n-1+n_x}\right)^{-1} \sum_{1\le s\le n} \binom{n-s+n_x}{n_x}$$
$$= \underbrace{\frac{n+n_x-1}{2(n-1)+n_x}}_{} \sum_{1\le s\le n} \binom{n-s+n_x}{n_x}$$

We can now compute a lower bound for $\hat{p}_{odd}(n, n_x)$.

$$\hat{p}_{odd}(n, n_x) \ge \left(\frac{n}{n+n_x+1}\right)\left(\frac{n+n_x-1}{2(n-1)+n_x}\right)$$
$$= \left(\frac{n}{2(n-1)+n_x}\right)\left(\frac{n+n_x-1}{n+n_x+1}\right)$$

□

1.9.3. Verifying Table 1

The proportion of coefficients that are of odd σ–order is displayed in Table 1. For instance, in a fourth–order perturbation ($n = 4$) with ten state variables ($n_x = 10$) the percent of coefficients of odd σ–order is 21.8%.

Let's exhaustively verify a couple entries in Table 1. Consider a first–order perturbation ($n = 1$) with one state variable ($n_x = 1$), the following coefficients need to be computed: $\mathbf{g}_{x_1}$, $\mathbf{h}_{x_1}$, $\mathbf{g}_\sigma$, and $\mathbf{h}_\sigma$. The first two are of the zeroth σ–order, and the latter two are of the first σ–order. In other words, half of the coefficients are of odd σ–order, which corresponds to the first entry of ".5" in Table 1.

Consider a second order perturbation ($n = 2$) with two state variables ($n_x = 2$). The coefficients that need to be computed are: $\boldsymbol{g}_{x_1}$, $\boldsymbol{g}_{x_2}$, $\boldsymbol{g}_{x_1 x_2}$, $\boldsymbol{g}_{x_1^2}$, $\boldsymbol{g}_{x_2^2}$, $\boldsymbol{g}_\sigma$, $\boldsymbol{g}_{x_1\sigma}$, $\boldsymbol{g}_{x_2\sigma}$, and $\boldsymbol{g}_{\sigma^2}$ (symmetrically for $\boldsymbol{h}$). Three of the nine, or $1/3 = .333$, are of the first σ–order. Notice that this does not depend on the number of control variables n_y.

All of the entries in Table 1 are less than 1/2. This is because there are more coefficients of order σ^s than σ^{s+1}. That is, there are more coefficients of order σ^0 than σ^1, and of order σ^2 than $\sigma^3 \ldots$ Hence, there are more even σ–order equations than odd. Theorem 1 does not apply to the (deterministic) coefficients of order σ^0, but does substantially reduce the number of other (stochastic) coefficients of higher σ–order.

CHAPTER 2

Redefining Information and Efficiency to Understand Technical Analysis

2.1. Introduction

The efficient market hypothesis (EMH), as established by Fama (1970, 1991), precludes any rationalization of technical analysis—despite being the dominant form of trading behavior at short horizons (Allen and Taylor, 1990; Cheung and Chinn, 2001; Gehrig and Menkhoff, 2006; Menkhoff, 2010). Until such a juxtaposition between theory and practice is resolved, economists cannot be confident in their understanding of markets, from policy and trading advice to the informational behavior of prices. The question is what information does public data contain that only technical analysts can effectively aggregate? Our answer is to (1) redefine information and (2) the EMH so that private information can be learned from public data and (3) illustrate why technical analysts are needed to aggregate "variably diffuse information."

First, the way information is defined in economics via probability theory makes no distinction between observing data and running every possible analysis. That is, there is no direct way to condition on pieces of information without also conditioning on all their exponentially many interactions. While useful for probabilistic foundations, this narrow definition fails to recognize that analyzing interactions is costly. Traders might reasonably know pieces of information without fully understanding their interactions. Rather than framing this as just bounded rationality, information should be defined broadly to include analysis.

Without changing any rules of probability spaces, we generalize how information is defined via σ-fields to *families* of events.[26] The point is that families do not have to be closed under intersection, allowing pieces of information to be conditioned on without their many interactions. While there is no obvious way to make this work probabilistically, it readily

[26]See Appendix 2.7 for a review of σ-fields and families using examples in Section 2.2.

follows when information is defined as knowledge about optimization solutions. Specifically, conditioning on a family means knowing how to optimize over all functions at least as *coarse*. In finite sample spaces, coarseness naturally extends and is consistent with the probabilistic definition of conditioning on σ-fields.

By viewing analysis as a form of information, it becomes apparent just how strong even the weakest version of the EMH really is. For markets to have no *excess return predictability*, all the interactions contained in troves of public time series data must effectively be common knowledge. Given the complexity of markets, there is little reason to expect this approximation holds *a priori* even with significant incentives. Indeed, the sheer number of traders analyzing public data to predict excess returns cannot be rationalized if interactions are so well understood.

Our basic version of the efficient market hypothesis (EMH*) then says that prices are as they would be if all private analyses of public data were common knowledge. This mathematically formalizes what information prices do not reflect in a way that could only be verbalized previously. Of course, analyzing data is costly, so prices only partially reflect public data to the extent traders have been incentivized to analyze it. As analytical methods improve over time, it is not particularly surprising under EMH* how many profitable trading strategies are found retroactively in the finance literature that were already obsolete (Park and Irwin, 2007).

While efficiency hypotheses are conceptually useful, there are many reasons why markets might not fully aggregate private analyses. For example, traders must still earn information rents as in Grossman and Stiglitz (1980). Rather, like the welfare theorems, sufficient conditions for hypotheses provide a starting point to understand what can go wrong, and digging into their nuances may deliver a flourishing theory. The first step is to appropriately define information, which has implications across economics, but most directly to the EMH and information aggregation. After generalizing σ-fields to families and deriving efficiency results, the question becomes what is the broadest way to define information?

Economics is about decision making, and the fundamental basis for it is knowledge about optimization solutions. This sort of knowledge might be familiar on a practical level because it includes everything naturally done when computation is limited, which is entirely the point—analysis is costly and constitutes its own form of information. We develop a mathematical structure around induction to describe how someone might be knowledgeable about optimization solutions and then make inferences. This gives a new perspective on behavioral questions and game theory, but it is likely too abstract and unconstrained for direct application.

Having belabored ways of defining analysis, we turn to why technical analysts are prevalent by explaining their unique informational role. How do patterns arise that other traders, such as value investors and liquidity providers, cannot effectively model? Our answer is that variation over information structures is hard to observe but affects price through its interaction with demand. Patterns then result between associated variables and price that seem economically spurious and cannot be readily derived.

More specifically, for any given asset and point in time, it is hard to know how many traders are informed about what. Even in the simplest model where traders draw from the same univariate signal distribution, there will inevitably be aggregate variation in the number of traders and their distributional moments. These interact with demand curvature and generate pricing errors to the extent traders do not account for them. While directly estimating such a latent process might be too difficult, theoretical models can at least illustrate how prices behave and what patterns to look for.

For example, Miller (1977) shows that unexpectedly disperse opinions cause higher prices through the winner's curse. Their effect is magnified by the relative wealth of potential traders to market capitalization, so economic policies and variables that increase this ratio can reduce price informativeness. This comparative static is notably the reverse of noise trading results stemming from Grossman (1976). Ultimately, many different sources of noise cumulatively determine the behavior of prices.

Only by modeling these sources of noise can traders better figure out which are significant in what markets and how to better account for them. Even if traders do not use well specified parametric models, theory can still guide what profitable strategies to look for. Dispersion and the winner's curse suggest that interactions between related variables need to be accounted for, such as: analyst variation, short interest, liquidity, and market depth. To what extent these interactions have been accounted for is then a question in the efficiency literature (Diether et al., 2002; Johnson, 2004; Sadka and Scherbina, 2007; Moeller et al., 2007).

More generally, we illustrate how random effects and distortions over information structures cause pricing errors with many different implications for trading and price informativeness. The complex interactions that arise in just our toy model belie how intractable they are to parametrically model in real markets. Technical analysts who broadly decrypt financial data may then have a comparative advantage finding these seemingly economically spurious interactions over conventional traders.

The remainder of this paper is organized as follows. Section 2.2 presents the EMH and how it is limited by information sets. Section 2.3 extends conditioning to families of events and gives the corresponding efficient market hypothesis (EMH*), for which Section 2.4 develops sufficient conditions. Section 2.5 illustrates how "variably diffuse information" can generate the kind of short-term economically spurious patterns that technical analysts look for, and Section 2.6 concludes.

2.2. The Efficient Market Hypothesis (EMH)

While many different versions of the EMH have been proposed, they all follow the same common formulation given by Jensen (1978):

> A market is efficient with respect to information set θ_t if it is impossible to make economic profits by trading on the basis of information set θ_t.

Each version of the EMH then corresponds to a specific choice of information set θ_t, which generally includes all past security prices—if not harder to access public data or even private information.

2.2.1. Specifying a Joint Hypothesis

The EMH is notably incomplete without a joint hypothesis that specifies exactly how information should translate into prices. Whether or not economic profits can be made by a given trader on the basis of information set θ_t depends on their preferences and priors. Implicitly, everything is being defined here with respect to some representative trader, so what are *reasonable* priors for them to have?

Given some canonical preferences, if it is to be said prices fully reflect information set θ_t, then we argue traders must be able to learn any sufficiently repeated pattern contained within. Patterns at shorter horizons with higher frequency will naturally be easier to demonstrate. For example, if there were a large Monday effect, traders with information sets containing all past security prices should know of it. Whereas, say the long-run effect of macroeconomic policies on fundamentals might require a much longer history to discern. This ability to learn patterns is a constraint on how degenerate and misspecified priors can be.

Unless we specify that knowing information entails having fully introspected the relevant priors, then knowledge itself can be an arbitrarily weak notion. At a minimum, we argue that fully introspected priors should not preclude the ability to learn patterns. While there remains a gray area as to what exactly constitutes a pattern or how much information a trader would need to learn one, it still conceptually narrows down what joint hypotheses pair meaningfully with the EMH.

2.2.2. The Problem with Information Sets

Having specified the relevant set of joint hypotheses, we can now try to explain the disconnect between trading behavior in practice and the EMH in theory. Our critique is that the very concept of information sets is too narrow: having pieces of information necessarily means being able to condition on all their interactions. There exists a whole possible spectrum of information by which any level of interactions could be known. We illustrate this limitation with Example 1 and Example 2.

Example 1 (Separability)**.** *Consider the following two random variables* $X_1 \in \{0, 1\}$ *and* $X_2 \in \{0, 1, 2\}$ *in the corresponding probability space* $(\Omega, 2^\Omega, \mathbb{P})$*.*

		X_2		
	Ω	0	1	2
X_1	0	ω_1	ω_2	ω_3
	1	ω_4	ω_5	ω_6

Table 8: Separability example.

In probability theory, information is defined with respect to σ-fields which describe what events can be conditioned on. (See Appendix 2.7 for a basic review of σ-fields in the context of these examples.) Notably, σ-fields are families of events that are closed under countable intersection and complementation. Observing an information set, e.g. a collection of random variables or family of events, necessarily means being able to condition on all of the events in the σ-field they generate. However, these closure properties make it so that pieces of information cannot be known without their interactions.

In the probability space given by Example 1, knowing X_1 and X_2 necessarily means being able to condition on all their interactions since $\sigma(X_1, X_2) = 2^\Omega$ is the total σ-field. Formally, let event E_i^j denote:

$$E_i^j = \{X_i = j\}$$

Note that any any singleton event $\{\omega_i\}$ in the sample space can be expressed as the intersection of (two) such events, e.g. $E_1^0 \cap E_2^0 = \{\omega_1\}$. By closure under intersections, there is no way someone could know the events corresponding to X_1 (E_1^0 and E_1^1) and X_2 (E_2^0, E_2^1, and E_2^2) without also knowing all their interactions (2^Ω).

Nor can separability be expressed via sufficient statistics because the generated sub σ-field would still need to be closed under intersection. One solution is to set up the information structure so that nothing can be learned from conditioning on interactions. However, this abstracts the content of knowledge away from the observed data to specifics of the information structure, e.g. priors and global games. This does not lend itself to an EMH that specifies what parts of publicly available financial data is and is not accounted for in market prices.

We want a general mathematical system where X_1 and X_2 can be known separately without their interactions. This means being able to condition on each of three and two events corresponding to X_1 and X_2 but not each of the six realizations in the sample space. Traders would effectively have five degrees of freedom instead of six. Economics, unlike probability theory, has to do with what people know rather than what could possibly be inferred. Rather than requiring information be closed under intersection, we suggest traders may know families of events.

While it would be trivial for traders to interact X_1 and X_2 in Example 1, the curse of dimensionality in large financial time series makes this distinction much more meaningful. If seemingly spurious interactions are prevalent in security prices, as trading behavior and Section 6 suggest, then there is little reason to expect each interaction would be known by some trader let alone by all. Traders may then be able to derive a great deal of private information from analyzing public data. The point is not just that analyzing data is costly, but rather that information sets are not specific enough to distinguish between what analyses are public, private, and unknown.

Example 2 (Curse of Dimensionality). *Consider the collection of random variables* $\{X_t^n\}_{t=1,\, n=1}^{T,\, N}$ *that can each take on* M *different possible values:* $X_t^n \in \{x_{t,1}^n, \ldots, x_{t,M}^n\}$. *The size of the sample space,* $|\Omega| = M^{TN}$, *is exponential in* T *and* N.

The point of Example 2 is merely that interactions can grow exponentially with the amount of data, i.e. the curse of dimensionality. It then becomes critical to distinguish between separately knowing pieces of information from their interactions, as the former grows linearly.

		X^n_{t+1}			
	Ω	0	1	$\dots$	$x^n_{t+1,m}$
X^n_t	0	ω_{i_1}	ω_{i_2}	$\dots$	ω_{i_m}
	1	$\omega_{i_{m+1}}$	$\omega_{i_{m+2}}$	$\dots$	$\omega_{i_{2m}}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$x^n_{t,m}$	$\omega_{i_{m^2-m+1}}$	$\omega_{i_{m^2-m+2}}$	$\dots$	$\omega_{i_{m^2}}$

Table 9: Cross section of table with T x N x M dimensions.

Otherwise, there is a bait-and-switch in our intuitions about the EMH. Any security price at a given point in time seems to clearly be public information and ought to be included by information set θ_t in the EMH. However, the σ-field generated by all past security prices taken together becomes large enough that there's no reason to expect *a priori* it will be fully accounted for.

Standard versions of the EMH are still conceptually and empirically useful for many reasons given by Fama (1991). However, we question here the underlying intuition for the EMH being a good approximation of markets. There is little reason to expect that all the information contained in security prices is approximately common knowledge and therefore reflected by prices. To the extent the EMH oversimplifies, it misleads traders about what strategies to look for.

2.3. Conditioning on Families of Events

The way information is defined via σ-fields makes no explicit distinction between simply observing pieces of information and analyzing their many interactions. This leads to a fallacious intuition where if each piece of information is publicly known, then the σ-field generated by their collection must be as well. Separability and curse of dimensionality are examples of what difficulties get incidentally shoved away. Our solution is to define information and conditioning with respect to families of events more generally, leading to alternate versions of the EMH.

2.3.1. Redefining Information

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the underlying probability space, then knowing a family $\mathcal{G} \subset \mathcal{F}$ is defined as being able to optimize over all functions with coarseness $\mathcal{G}$. A function has coarseness $\mathcal{G}$ if it can be written as the sum of indicator functions over events in $\mathcal{G}$, which Definition 1 formalizes. To avoid complications with infinite sums, we assume the sample space is finite, which could be as large as any bounded digitized data set.

Our following notion of *coarseness* extends the concept of measurability from σ-fields to families of events more generally.

Definition 1 (Coarseness)**.** *The set $F_{\mathcal{G}}$ of all functions with (finite) coarseness $\mathcal{G}$ is:*

$$F_{\mathcal{G}} \equiv \left\{ f(\omega) \;\middle|\; f(\omega) = \sum_{E \in \mathcal{G}} c_E \mathbb{1}_{\{\omega \in E\}} \text{ for some } \{c_E\}_{E \in \mathcal{G}} \text{ constants with respect to } \Omega \right\}$$

Analogously, $F_{\mathcal{G}}(y)$ is the set of all such functions where c_E can be a function of y.

Conditioning on a family $\mathcal{G}$ means being able to optimize over all functions with coarseness $\mathcal{G}$. Definition 2 characterizes a solution as being that which optimizes the unconditional expectation.

Definition 2 (Argmax)**.** *The set of solutions to a maximization problem conditional on a family $\mathcal{G}$ is given by:*

$$\begin{aligned} \underset{x \in X}{\arg\max}\; u(x \,|\, \mathcal{G}) &\equiv \underset{\hat{x} \in F_{\mathcal{G}},\ \hat{x}(\Omega) \subseteq X}{\arg\max}\; \mathbb{E}\left[u(\hat{x}(\omega))\right] \\ &\subseteq \{x^* : \Omega \to X\} \end{aligned}$$

Where the function u depends on the imperfectly observed realization ω.

Having defined arg max, all of the corresponding definitions for max/min and sup/inf readily follow.

Notably, if $\mathcal{G}$ is a σ-field, then conditioning turns into its standard probabilistic definition. The reason being that the smallest events in $\mathcal{G}$ generate a partition over Ω that can be separately optimized over. Thus, if $\mathcal{G}$ is a σ-field, then for any $\omega \in \Omega$, any optimal solution $x^*(\cdot \mid \mathcal{G})$ is necessarily such that:

$$x^*(\omega \mid \mathcal{G}) \in \arg\max_{x \in X} \mathbb{E}[u(x)|E_\omega], \text{ where } E_\omega = \bigcap_{E \in \mathcal{G},\, \omega \in E} E$$

Property 1. *Conditioning on families (Definition 2) is consistent with the standard probabilistic definition for conditioning on σ-fields and information sets.*

2.3.2. Generalizing the EMH

Having generalized the concept of information, we can now consider a much broader class of efficient market hypotheses than Jensen (1978). Each hypothesis here may correspond to a family $\mathcal{G}_t$ rather than an information set θ_t:

> A market is efficient with respect to family $\mathcal{G}_t$ if it is impossible to make economic profits by trading on the basis of family $\mathcal{G}_t$.

The critical distinction being that a family $\mathcal{G}_t$ may include pieces of information without all of their interactions. What interactions have been analyzed can readily be demarcated using families.

Specifically, if $\mathcal{G}_t$ is the union of all private analyses, $\mathcal{G}_t = \cup_i \mathcal{G}_t^i$ (where traders are indexed by i), then the EMH becomes:

Definition 3 (EMH*). *A market is informationally efficient if prices need not differ from the counterfactual where all private analyses of public data are common knowledge.*

To avoid the notion of economic profits because they depend on individual preferences, efficiency is defined here in terms of equilibria.

It should be emphasized that EMH* is no empirical substitute for standard versions of the

EMH that provide useful null hypotheses to test against. Rather, it challenges the underlying intuition for markets approximately aggregating all of the information contained in public data. The vastness of security price data combined with the curse of dimensionality leaves a potentially large gap between the information contained within and the union of all private analyses.

2.4. Sustaining Efficient Demand

The goal here is to find sufficient conditions under which an aggregate demand function D^* that leads to EMH* can be sustained in equilibrium. Our model is set up in such a way that the only question is whether traders with the relevant information have enough liquidity to construct D^*. If they can construct D^*, then because it is efficient, their preferences will fall into place such that it can be sustained in equilibrium.

Specifically, we consider a continuum of traders who are risk-neutral and liquidity constrained. Traders choose demand $x_1, \ldots, x_K$ in dollars for K different risky assets and save the rest of their wealth in a risk-free storage technology. Traders are able to use their liquidity to short trade an asset and receive a rate of return equal to that of the storage technology minus the asset. The realized total value of these assets $v_1, \ldots, v_K$ are positive random variables measurable with respect to the underlying finite probability space $(\Omega, 2^\Omega, \mathbb{P})$.

Given an aggregate demand function D^*, since the probability space is finite, there are finitely many types of traders $i \in \{1, \ldots, I\}$ characterized by their information, i.e. family of events $\mathcal{G}_i \subseteq 2^\Omega$ (which may depend on the prices that result from D^*). Because traders are small and risk neutral, if D^* has coarseness $\mathcal{G}^* = \cup_i \mathcal{G}_i$ and leads to EMH*, then a trader of type i is indifferent between all their possible trading strategies if $\mathcal{G}_i \subseteq \mathcal{G}^*$. Thus, we just need to show that the traders are capable together of constructing D^* given each information type has some liquidity l_i.

Let aggregate demand D^* have coarseness $\mathcal{G}^*$, then it can be decomposed into:

$$D^*(\omega) = \sum_{E \in \mathcal{G}^*} c_E^* \mathbb{1}_{\{\omega \in E\}}$$

The coefficients $|c_E^*|$ give an upper bound on how much liquidity needs to be"reserved" towards the event if it happens. That is, we need to be able to match trader types to events in $\mathcal{G}^*$ in sufficient measure to construct all the coefficients c_E^*.

This turns into the following bipartite matching problem where the inflow to trader type i is their liquidity l_i, and the required outflow from each event is c_E^*. Sufficient conditions for there being a solution to this matching problem are given by Hall's marriage theorem.

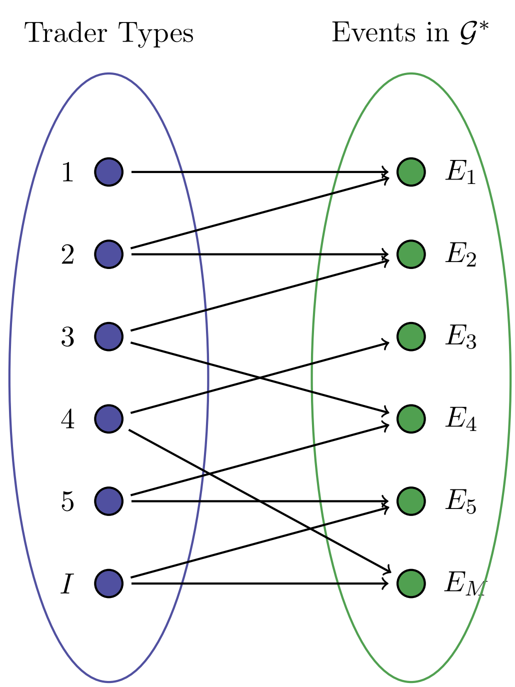

Figure 2: Bipartite flow matching problem with *events* in $\mathcal{G}^*$.

Result 1. *An aggregate demand function D^* with coarseness $\mathcal{G}^* = \cup_i \mathcal{G}_i$ that leads to equal returns across assets (conditional on $\mathcal{G}^*$) can be supported in equilibrium if for all families*

$\hat{\mathcal{G}} \subseteq \mathcal{G}^*$:

$$\sum_{E \in \hat{\mathcal{G}}} |c_E^*| \leq \sum_{i \,:\, \mathcal{G}_i \cap \hat{\mathcal{G}} \neq \emptyset} l_i$$

Where c_E^* *corresponds to the constant multiplying* $\mathbb{1}_{\{\omega \in E\}}$ *in the decomposition of* D^*.

Proof. Hall's marriage theorem and the triangle inequality (see Appendix 2.8). □

Intuitively, these sufficient conditions are lower bounds for the liquidity different subsets of traders need to have. Enough traders with the right information are needed to construct D^*. One glaring limitation of our result is that traders need to "reserve" liquidity for an event regardless of whether it is realized or not. This can be partially ameliorated by combining mutually disjoint events as follows.

Decomposition 1. *Let* D^* *be expressed as:*

$$D^*(\omega) = \sum_k c_k^* H_k$$

Where H_k *has coarseness* $\mathcal{H}_k \subseteq \mathcal{G}^*$ *a family of disjoint events and* $\|H_k\|_1 \leq 1$. *Note that* $\mathcal{H}_k$ *could still a singleton event.*

Theorem 2 (Sustaining Efficiency)**.** *An aggregate demand function* D^* *with coarseness* $\mathcal{G}^* = \cup_i \mathcal{G}_i$ *that leads to equal returns across assets (conditional on* $\mathcal{G}^*$*) can be supported in equilibrium if for all subsets* $\mathcal{K} \subseteq \{1, \ldots, K\}$:

$$\sum_{k \in \mathcal{K}} |c_k^*| \leq \sum_{\{i \,:\, \exists k \in \mathcal{K} \, s.t. \, \mathcal{H}_k \subseteq \mathcal{G}_i\}} l_i$$

Where c_k^* *corresponds to the coefficient of* H_k *in Decomposition 1 of* D^*.

Proof. Hall's marriage theorem and the triangle inequality (see Appendix 2.8). □

Much like the welfare theorems, the point here is not to say that markets are informationally

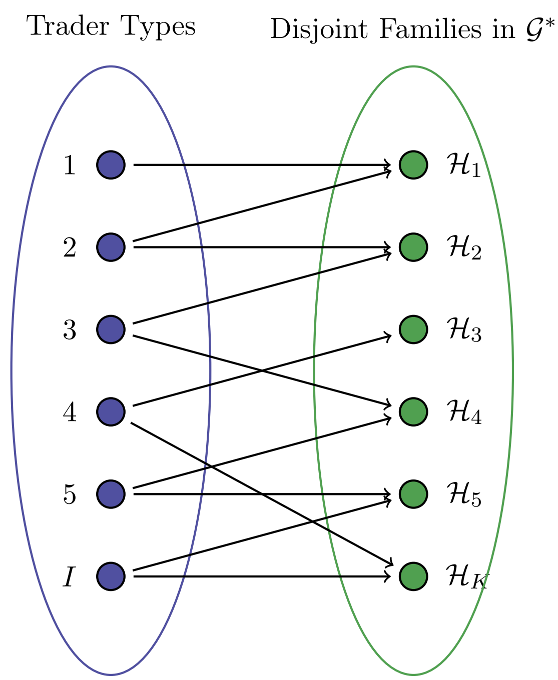

Figure 3: Bipartite flow matching problem with *disjoint families* in $\mathcal{G}^*$.

efficient. Rather, it is to provide a framework for understanding why they may be more or less efficient. To the extent our sufficient conditions do not hold, markets might not be able to aggregate private analyses.

2.5. Variably Diffuse Information

Defining analysis as a form of information helps rectify our intuitions about how difficult it is for markets to aggregate public data. Our EMH* and aggregation results help flesh out the concept of analysis, but they do not explain the need for technical analysts in particular. The next question is, how do patterns arise that seem economically spurious and cannot be well accounted for by value investors or liquidity providers?

In standard linear rational expectations equilibrium models (Grossman and Stiglitz, 1980; Hellwig, 1980; Kyle, 1985, 1989; Rostek and Weretka, 2012), value investors mainly counteract additive demand (and supply) shocks known as noise trading (Goldstein and Yang, 2015; Lambert et al., 2018). As the number of value investors grows large, prices become fully revealing in many of these models. However, there are many other potential sources of noise that value investors may not be able to aggregate away, e.g. the winner's curse in Miller (1977).

This section considers homogenous traders who draw signals for the value of an asset from some aggregate distribution with *variable moments*. While hard to economically interpret, this variation in how information is spread out across traders will interact with demand curvature and generate pricing errors. To the extent traders have not already accounted for such errors, there will be seemingly spurious patterns between related variables. For example, short interest and analyst ratings relate to distributional moments, and macroeconomic variables affect demand curvature through liquidity and risk preferences.

Using as simple of a model as possible, we illustrate many different ways that information could be variably diffuse and the resulting price behaviors. Appendix D grounds this model in the standard rational expectations literature with simplifying constraints on trading behavior. For exposition, we present the simplified model here on its own.

2.5.1. Toy Model Setup

Consider a continuum of homogenous traders buying one divisible asset. Each trader receives a private signal $s \sim N(r, \sigma_s^2)$ for the log rate of return r and then chooses demand $D(s)$ for the asset in terms of the numeraire. Notably, the input to the trader's problem here is unidimensional. Rather than observing both the price and an estimate for asset value, traders divide the two and get a single estimate s for log return r.

Supply of the asset in terms of the numeraire is simply its price, which is inversely related to log return r. Given some demand shock $\epsilon \sim N(0, \sigma_\epsilon^2)$, an equilibrium is characterized by demand D^* and log return $r^* = R^*(\epsilon)$ satisfying market clearing:

$$w(\epsilon, r^*) \equiv \underbrace{\log\left(\mathbb{E}_s\left[D^*(s) \,|\, r^*\right]\right)}_{\text{Log Aggregate Demand}} + \epsilon \; + \; r^* = 0 \tag{2.1}$$

Intuitively, higher demand leads to higher prices and lower rates of return. While the market clearing condition here is taken as given, it derives from a rational expectations equilibrium model in Appendix D.

Lastly, the equilibrium demand function must be a best-reply:

$$D^*(s) \equiv \arg\max_{x \subseteq X} \mathbb{E}_r \left[u(x, r) \mid s, R^* \right] \tag{2.2}$$

Where u is the traders ex-post utility given demand x (within budget constraint X) and rate of return r.

Table 10: Invariant equilibrium objects.

$R(\epsilon)$	Log return
$D(s)$	Demand

2.5.2. General Equilibrium Properties

This unidimensional model has the following well-behaved equilibrium properties. Since $s \mid r$ satisfies the monotone likelihood ratio property (MLRP) due to normality, demand is increasing for any return function (assuming $u_{xr} > 0$, single-crossing).

Property 2. *$D^*(s \mid R)$ is strictly increasing in s for any nonconstant $R(\epsilon)$.*

Because aggregate demand is increasing in any equilibrium, market clearing is more sensitive to returns than demand shocks, $w_r \geq w_\epsilon = 1$.

Corollary 4. *Any equilibrium $R^*(\epsilon)$ is strictly decreasing in ϵ.*

Corollary 5. *The distribution of $R^*(\epsilon)$ is a weak contraction of ϵ.*

For a large class of distributions (MLRP) and utility functions (single-crossing), our model only admits monotone equilibria that are analytically useful. Corollary 5 adds that returns will always be less noisy than the underlying demand shocks.

2.5.3. The Winner's Curse Extended

The logic of the winner's curse in Miller (1977) is that, when there are many traders who cannot short trade, then unaccountedly disperse signals lead to high prices and subpar returns. We generalize it here as the interaction between variance and demand curvature.

Observation 1. *If equilibrium demand D^* is convex/concave, then unaccountedly high signal variance σ_s^2 leads to greater/lower demand and therefore lower/greater returns.*

Estimating such a latent interaction is difficult especially if traders are not even aware of it and only focus on fundamentals. Our point is that a whole host of economic variables could relate to either this signal variance or demand curvature, and the latent as well as spurious nature of these interactions give a comparative advantage to general pattern recognition over structural economic models. This thereby creates a potential niche for technical analysts who profit off these shorter fluctuations rather than longer term fundamentals.

Another variable that could have unexpectedly high variation are the demand shocks. How are returns effected (in expectation) during periods of unexpectedly high demand volatility? The answer is somewhat analogous but notably depends on the curvature of aggregate demand instead of individual demand.

Observation 2. *If equilibrium log aggregate demand* $\log\left(\mathbb{E}_s\left[D^*(s) \,|\, r^*\right]\right)$ *is convex (concave) with respect to return r^*, then unaccountedly high demand shock variance σ_ϵ^2 leads to lower (greater) returns in expectation.*

This rigorously follows from the implicit function theorem. Intuitively, if log aggregate demand were convex, then large negative demand shocks have diminishing effects on returns while large positive demand shocks have increasing effects.

The complex interactions in this simple unidimensional model belie how difficult it is for markets to aggregate variably diffuse information and the need for non-traditional traders. Traders caring in part about log returns means curvature is unavoidable here.

Observation 3. *Individual demand* $D^*(s)$ *and log aggregate demand* $\log\left(\mathbb{E}_s\left[D^*(s)\,|\,r^*\right]\right)$ *cannot both be linear.*

While linear rational expectations equilibrium models are analytically tractable, given liquidity constraints and the variety of risk preferences that traders have, there is little reason to expect linear demand in practice.

However, without linearity, problems aggregating variably diffuse information arise rather generally beyond variances.

Observation 4. *Variation in moments of* σ_s^2 *and* σ_ϵ^2 *interact with the corresponding Taylor series orders of* $D^*(s)$ *and* $\log\left(\mathbb{E}_s\left[D^*(s)\,|\,r^*\right]\right)$ *respectively to generate pricing errors.*

2.5.4. Cutoff Demand Example

For a concrete example, we suppose investors are risk neutral (over log returns) but liquidity constrained. Then, from property 2, individual demand is without loss of generality the cutoff rule:

$$D^*(s\,|\,R) = \begin{cases} L & s \geq s^* \\ 0 & s < s^* \end{cases}$$

Aggregate demand, where Φ is the normal cdf, is:

$$\mathbb{E}_s\left[D(s)\,|\,\epsilon\right] = L \cdot \Phi\left(-z^*(\epsilon)\right)$$

$$z^*(\epsilon) \equiv \frac{s^* - R^*(\epsilon)}{\sigma_s}$$

And, market clearing becomes:

$$\log\left(\Phi(-z^*(\epsilon))\right) + \log(L) + \epsilon + R^*(\epsilon) = 0$$

We can understand price informativeness through the magnitude of the slope of R^*, which is given by the implicit function theorem as:

$$R_\epsilon^*(\epsilon) = \frac{-1}{H(z^*(\epsilon)) + 1}$$

Where H is the standard normal hazard function, which increases in z^*. The flatter $R^*(\epsilon)$ is the more informative prices are.

Property 3. *There is a unique equilibrium and $z^*(\epsilon)$ strictly increases in liquidity L.*

Then, $|R_\epsilon^*|$ is decreasing in liquidity and prices become more informative with liquidity.

Corollary 6. *For $L_1 < L_2$, $R^*(\epsilon \,|\, L_1)$ is a weaker contraction of ϵ than $R^*(\epsilon \,|\, L_2)$.*

This result follows generally for non-normal distributions if the hazard function H is increasing.

Because the normal hazard function grows arbitrarily large, we get that prices become arbitrarily informative with liquidity, or equivalently the measure of traders here, which aligns with the literature.

Corollary 7. $\lim_{L\to\infty} z^*(\epsilon \,|\, L) = \infty$, *and* $\lim_{L\to\infty} R^*(\epsilon \,|\, L) = \bar{r}$ *the implied outside rate of return where* $u_x(0, \bar{r}) = 0$.

However, note that in this example, log aggregate demand $\log\big(\Phi(-z^*(\epsilon))\big)$ is a concave function of r (because the normal hazard function is increasing). Thus, it follows from Remark 2 that markets with unexpectedly high demand shock variance σ_ϵ^2 are underpriced and give greater returns.

Observation 5. *In this cutoff demand example, assets give greater returns in expectation when the demand shock variance σ_ϵ^2 is unexpectedly high.*

2.6. Conclusion

The purpose of this paper is to reconcile classical intuitions about efficient markets with the prevalence of technical analysts in practice. Our first step is to define analysis as a form of information, thereby distinguishing individual pieces of information from their many interactions. This new information structure allows for bounded rationality models where analysis is costly, such as correlation neglect. It then becomes apparent how strong the EMH is by suggesting markets account for all the information contained in public data—even analyses no one has conducted.

While the EMH is an empirically useful null hypothesis, a more conceptually sound version is that markets aggregate all private analyses. The question then becomes why might markets be more or less informationally efficient? To this end, we develop sufficient conditions for sustaining efficient aggregate demand in equilibrium—giving a framework for what could go wrong. These rudimentary results further serve as examples of how to do theoretical work with families of events as a new form of information.

Lastly, to reconcile theory and practice, we need to explain the unique informational role technical analysts play. Rather than directly modeling fundamentals or liquidity, they focus on broad statistical pattern recognition at short horizons. Section 5 illustrates how latent interactions between distributional variation and demand curvature can generate these broad short term patterns that seem economically spurious. This suggests technical analysts have a comparative advantage aggregating variably diffuse information, which gives empirical implications for what markets they are prevalent in.

2.7. Appendix: Review of σ-Fields

A σ-field $\mathcal{F}$ is any family of events, i.e. $\mathcal{F} \subseteq 2^{\Omega}$ where Ω is the sample space, satisfying the three properties that it:

1. Is closed under complements, if $E \in \mathcal{F}$ then $E^c \in \mathcal{F}$.

2. Contains the sample space, $\Omega \in \mathcal{F}$. (This implies that $\emptyset \in \mathcal{F}$ by closure under complements.)

3. Is closed under countable intersection, if $E_1, E_2, \ldots \in \mathcal{F}$ then $\cap_{i=1}^{\infty} E_i \in \mathcal{F}$. (This implies closure under countable unions because of closure under complements.)

A probability space is defined by a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space and $\mathbb{P}$ is the probability measure defined over the σ-field $\mathcal{F}$. When the sample space is finite, then $\mathcal{F}$ is often taken to be the "total σ-field," $\mathcal{F} = 2^{\Omega}$. Much of the impetus for σ-fields is that complications arise when the sample space is larger, e.g. $\Omega = \mathbb{R}$, where $\mathbb{P}$ cannot be defined over the power set without violating axioms of probability.

Information is always characterized by the sub σ-field it generates in the underlying probability space. Observing some family of events $\{E_i\}_{i \in \mathcal{I}}$ corresponds to observing all events in the generated σ-field:

$$\sigma\big(\{E_i\}_{i \in \mathcal{I}}\big) \equiv \bigcap_{\substack{\hat{\mathcal{F}} \text{ is a } \sigma\text{-field} \\ \{E_i\}_{i \in \mathcal{I}} \subseteq \hat{\mathcal{F}}}} \hat{\mathcal{F}}$$

That is, the information someone has is the closure under the three σ-field properties of the events they observe. If someone observes a random variable, then they observe events corresponding to each possible value of that variable. The knowledge an information set describes is the σ-field generated by it.

Our paper makes no contribution to probability theory which describes what can be inferred from pieces of information. However, unlike probability, economics needs a broader definition of information that recognizes the cost of analyzing interactions. We argue the way to do this is by letting any family $\mathcal{G} \subseteq \mathcal{F}$ characterize knowledge. The following examples illustrate the usefulness of this family notion.

2.7.1. Detailing the Separability Problem in Example 1

Consider the two random variables X_1 and X_2 in probability space $(\Omega, 2^\Omega, \mathbb{P})$ from Table 1. Knowing X_1 corresponds to observing the two events:

$$E_1^0 = \{X_1 = 0\} = \{\omega_1, \omega_2, \omega_3\}$$

$$E_1^1 = \{X_1 = 1\} = \{\omega_4, \omega_5, \omega_6\}$$

And, knowing X_2 corresponds to observing the three events:

$$E_2^0 = \{X_2 = 0\} = \{\omega_1, \omega_4\}$$

$$E_2^1 = \{X_2 = 1\} = \{\omega_2, \omega_5\}$$

$$E_2^2 = \{X_2 = 2\} = \{\omega_3, \omega_6\}$$

Any event $W_i = \{\omega_i\}$ can be written as the intersection of E_1^0 or E_1^1 with one of E_2^0, E_2^1, or E_2^2. For example, $W_1 = \{\omega_1\} = E_1^0 \cap E_2^0$, which is the event that both X_1 and X_2 are zero. Since any singleton event can be generated this way, and σ-fields are closed under countable union:

$$\sigma(E_0^1, E_1^1, E_0^2, E_1^2, E_2^2) = 2^\Omega$$

That is, there is no way in the given probability space to describe separately knowing X_1 and X_2 but not their interactions. Anyone who observes the five events E_0^1, E_1^1, E_0^2, E_1^2, and E_2^2 must necessarily also know how to condition on each of the six possible realizations in the sample space Ω.

Notably, sufficient statistics cannot solve this separability problem since they still must generate a sub σ-field of 2^Ω, which cannot be just the five aforementioned events. One solution is to define a larger information structure, e.g. priors or a global game, whereby the interaction between the two variables is meaningless. Even though their interaction can be conditioned on, the two variables are effectively separable. However, this does not lend itself to an efficient market hypothesis that directly specifies what data is and is not accounted for.

By using families instead of σ-fields to define information, we are able to directly specify notions such as separability in terms of conditioning on data. Traders are able to separately know X_1 and X_2 without their interactions. While the difference between conditioning on five events versus each of the six realizations in the sample space is relatively small, and analyzing their interactions would be trivial, the nuance becomes critical in large data sets.

2.8. Appendix: Sustaining Efficient Demand

This appendix proves Theorem 2 and therefore the special case of Result 1.

Proof. By assumption, Hall's marriage theorem applies to the bipartite flow matching problem in Figure 2. Hence, there exists a flow $\{f_{k,i}\}_{k=1,i=1}^{K,I}$ such that:

$$\sum_{i=1}^{I} f_{k,i} = c_k^* \quad \text{and} \quad \sum_{k=1}^{K} f_{k,i} \leq l_i$$

Let all traders of information type i have the demand function:

$$D_i = \frac{1}{l_i} \sum_{k=1}^{K} f_{k,i} H_k$$

By construction, each trader's liquidity constrain is satisfied:

$$\begin{aligned} \|D_i(\omega)\|_1 &= \|\frac{1}{l_i} \sum_{k=1}^{K} f_{k,i} H_k\|_1 \\ &\leq \frac{1}{l_i} \sum_{k=1}^{K} f_{k,i} \|H_k\|_1 \\ &\leq 1 \end{aligned}$$

And, the traders together aggregate to the efficient demand function D^*.

$$\begin{aligned}\sum_{i=1}^{I} l_i D_i &= \sum_{i=1}^{I}\sum_{k=1}^{K} f_{k,i} H_k \\ &= \sum_{k=1}^{K} c_k^* H_k \\ &= D^*\end{aligned}$$

Since D^* results in assets having equal returns conditional on $\mathcal{G}^* = \cup_i \mathcal{G}_i$, small risk neutral traders are indifferent between all of their possible demand functions. □

2.9. Appendix: Variably Diffuse Information

The point of this section is to

2.9.1. Setup

One divisible asset of common value v is traded on a financial market with a continuum of investors. The total measure of investors is m and each observes a private signal $\hat{s}$. Investors can buy the asset at an endogenously determined price p and then realize a log rate of return $r = \log(v/p)$. The ex-post utility from investing x dollars is given by a vNM utility function $u(x, r)$ with basic regularity conditions.

Table 11: Rational expectations variables.

v	Value
p	Price
r	Log return
m	Measure of investors
$\hat{s}$	Private signal
x	Investment
ϵ	Mismatch

The two exogenous variables in the model are value v and *mismatch* $\epsilon \equiv \log(m/v)$:

$$v \sim \log N(\mu_v, \sigma_v^2)$$

$$\epsilon \equiv \log(m/v) \sim N(\mu_\epsilon, \sigma_\epsilon^2)$$

These variables are independent, which is an economic assumption that investors match with assets in proportion to value.

The distribution of private signals is centered around value:

$$\hat{s}|v \sim \log N(\log(v), \sigma_s^2)$$

Since investors lie on a continuum, this is the realized aggregate distribution of all private signals. Investors are otherwise homogenous and so their behavior is modeled with a representative agent. Everything about the model is common knowledge.

For a given demand function $\hat{D}(\hat{s}, p)$, the market clears if:

$$m \cdot \mathbb{E}_{\hat{s}}\left[\hat{D}(\hat{s}, p) \,\middle|\, v\right] = p$$

$$\Leftrightarrow \mathbb{E}_{\hat{s}}\left[\hat{D}(\hat{s}, p) \,\middle|\, v\right] = \frac{p}{v} \cdot \frac{v}{m}$$

$$\Leftrightarrow \log\left(\mathbb{E}_{\hat{s}}\left[\hat{D}(\hat{s}, p) \,\middle|\, v\right]\right) + r + \epsilon = 0$$

From this equation we define an equilibrium as follows.

Definition 4. *An equilibrium is characterized by a log return function $\hat{R}^*(v, \epsilon)$ that clears the market for all v and ϵ:*

$$\underbrace{\log\left(\mathbb{E}_{\hat{s}}\left[\hat{D}^*\left(\hat{s}, p \,\middle|\, \hat{R}^*\right) \,\middle|\, v\right]\right)}_{\textit{Log Aggregate Demand}} + \hat{R}^*(v, \epsilon) + \epsilon = 0 \tag{2.3}$$

Where $p = ve^{-\hat{R}^*(v,\epsilon)}$, and the expected utility maximizing level of investment is:

$$\hat{D}^*\left(\hat{s}, p \,\middle|\, \hat{R}\right) \equiv \arg\max_{x \in \mathbb{R}} \mathbb{E}_r\left[u(x, r) \,\middle|\, \hat{s}, p, \hat{R}\right] \tag{2.4}$$

Table 12: General equilibrium objects.

$\hat{R}(v, \epsilon)$	Price
$\hat{D}(\hat{s}, p)$	Demand

2.9.2. Scale Invariance

What does the equilibrium demand function look like? Intuitively, the most relevant piece of information to investors is the ratio of their signal to the price:

$$s \equiv \log(\hat{s}/p)$$

Along this line of thought, Section 2.9.3 supposes investors can only condition on their normalized signal s, thus demand is given by $D(s)$. Then, Claim 5 shows market clearing returns can be selected independent of value as $R(\epsilon)$. This leads to a simple definition of equilibria where price and demand are univariate functions.

Using this simplified model, Section 2.5.2 proves various equilibrium properties, e.g. $R^*(\epsilon)$ is a decreasing contraction. These properties then let us justify conditioning on normalized signals as rational inattention as well as argue it approximates equilibria of the general model.

2.9.3. Equilibria

If investors only condition on s, the following claim proves that: under any demand function for which there exist market clearing returns, there exist market clearing returns that do not depend on value.

Claim 5. *Given demand $D(s)$, if there exist returns $\hat{R}(v, \epsilon)$ that clear the market, then there exist returns $R(\epsilon)$ that clear the market as well.*

Proof. Fix $\hat{v}$ and let $R(\epsilon) \equiv R(\hat{v}, \epsilon)$, then $R(\epsilon)$ clears the market for $\hat{v}$ and all ϵ by construction. Now, normalized signals $s = \log(\hat{s}/p)$, where $p = ve^{-R(\epsilon)}$, have the distribution:

$$s = \log(\hat{s}/v) + R(\epsilon) \sim N\big(R(\epsilon), \sigma_s^2\big)$$

The market clearing equation is then constant in v:

$$\Leftrightarrow \log\big(\mathbb{E}_s\left[D(s) \,|\, \epsilon\right]\big) + R(\epsilon) + \epsilon = 0$$

Hence, $R(\epsilon)$ clears the market for all v and ϵ. □

If investors only condition on s, and markets select clearing returns that do not depend value, then an equilibrium is defined as follows.

Definition 5. *A "scale invariant equilibrium" is characterized by a log return function $R^*(\epsilon)$ that clears the market for all mismatch ϵ:*

$$\log\Big(\mathbb{E}_s\left[D^*(s \,|\, R^*) \,\big|\, \epsilon\right]\Big) + R^*(\epsilon) + \epsilon = 0 \tag{2.5}$$

Where investors maximize their expected utility:

$$D^*(s \,|\, R) \equiv \arg\max_x \mathbb{E}_r\left[u(x, r) \,|\, s, R\right] \tag{2.6}$$

With the signal distribution $s \sim N\big(R(\epsilon), \sigma_s^2\big)$.

Table 13: Scale invariant functions.

$R(\epsilon)$	Log return
$D(s)$	Demand

Since $\log(p) = \log(v) + R^*(\epsilon)$ where v and $R^*(\epsilon)$ are independent, Corollary 5 lets us bound how informative the price level is about returns.

Corollary 8. *$Corr\big(\log(p), R^*(\epsilon)\big) = \sigma_r/\sqrt{\sigma_v^2 + \sigma_r^2} < \sigma_\epsilon/\sqrt{\sigma_v^2 + \sigma_\epsilon^2}$.*

The price level only conveys information about returns through the prior over values. Thus, it is arbitrarily uninformative for large $\sigma_v/\sigma_r > \sigma_v/\sigma_\epsilon$. This holds for both informed and uninformed investors since the signal distribution is independent of price level (conditional on returns).

Therefore, scale invariant equilibria can be supported by *rational inattention* to price level if the prior over values is sufficiently flat. As this prior represents the beliefs of investors before observing a signal or the price, it can be interpreted as the distribution of all values over some relevant broad class of assets. Thus, we argue flat priors are plausible, and that it would be difficult for investors to properly estimate and generally account for σ_v.

To the extent investors do account for σ_v, we expect the price level to become even less informative about returns. Thus, an equilibrium in this simplified model ought to closely approximate that in the general model. (Note that we do not expect to see the sort of unravelling that can happen with priors in coordination games.)

BIBLIOGRAPHY

H. Allen and M. Taylor. Charts, noise and fundamentals in the london foreign exchange market. *The Economic Journal*, 100(400):49–59, 1990. ISSN 00130133, 14680297. URL http://www.jstor.org/stable/2234183.

M. Andreasen. On the effects of rare disasters and uncertainty shocks for risk premia in non-linear DSGE models. *Review of Economic Dynamics*, 15(3):295 – 316, 2012. ISSN 1094-2025.

B. Aruoba, J. Fernandez-Villaverde, and J. Rubio-Ramirez. Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control*, 30(12): 2477–2508, 2006.

D. Caldara, J. Fernandez-Villaverde, J. Rubio-Ramirez, and W. Yao. Computing DSGE models with recursive preferences and stochastic volatility. *Review of Economic Dynamics*, 15:188 – 206, 2012. ISSN 1094-2025.

Y.-W. Cheung and M. Chinn. Currency traders and exchange rate dynamics: a survey of the us market. *Journal of International Money and Finance*, 20(4):439–471, 2001. ISSN 0261-5606. doi: https://doi.org/10.1016/S0261-5606(01)00002-X.

O. de Groot. Solving asset pricing models with stochastic volatility. *Journal of Economic Dynamics and Control*, 52:308 – 321, 2015. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2015.01.001.

O. de Groot. What order? Perturbation methods for stochastic volatility asset pricing and business cycle models. *University of St Andrews Manuscript*, 2016.

K. Diether, C. Malloy, and A. Scherbina. Differences of opinion and the cross section of stock returns. *The Journal of Finance*, 57(5):2113–2141, 2002. doi: https://doi.org/10.1111/0022-1082.00490.

E. Fama. Efficient capital markets: A review of theory and empirical work*. *The Journal of Finance*, 25(2):383–417, 1970. doi: 10.1111/j.1540-6261.1970.tb00518.x.

E. Fama. Efficient capital markets: Ii. *The Journal of Finance*, 46(5):1575–1617, 1991. doi: https://doi.org/10.1111/j.1540-6261.1991.tb04636.x.

J. Fernandez-Villaverde, P. Guerron-Quintana, J. Rubio-Ramirez, and M. Uribe. Risk Matters: The Real Effects of Volatility Shocks. *The American Economic Review*, 101(6):2530 – 2561, 2011.

J. Fernandez-Villaverde, J. Rubio-Ramirez, and F. Schorfheide. Chapter 9 - solution and estimation methods for dsge models. In J. B. Taylor and H. Uhlig, editors, *Handbook of*

Macroeconomics, volume 2, pages 527 – 724. Elsevier, 2016.

T. Gehrig and L. Menkhoff. Extended evidence on the use of technical analysis in foreign exchange. *International Journal of Finance & Economics*, 11(4):327–338, 2006. doi: https://doi.org/10.1002/ijfe.301.

I. Goldstein and L. Yang. Information diversity and complementarities in trading and information acquisition. *The Journal of Finance*, 70(4):1723–1765, 2015. doi: 10.1111/jofi.12226.

S. Grossman. On the efficiency of competitive stock markets where trades have diverse information. *The Journal of Finance*, 31(2):573–585, 1976. ISSN 00221082. URL http://www.jstor.org/stable/2326627.

S. Grossman and J. Stiglitz. On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3):393–408, 1980. ISSN 00028282. URL http://www.jstor.org/stable/1805228.

M. Hellwig. On the aggregation of information in competitive markets. *Journal of Economic Theory*, 22(3):477 – 498, 1980. ISSN 0022-0531. doi: https://doi.org/10.1016/0022-0531(80)90056-3.

M. Jensen. Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2):95–101, 1978. ISSN 0304-405X. doi: https://doi.org/10.1016/0304-405X(78)90025-9.

H.-H. Jin and K. Judd. Perturbation methods for general dynamic stochastic models. *Manuscript, Stanford University*, 2002.

T. Johnson. Forecast dispersion and the cross section of expected returns. *The Journal of Finance*, 59(5):1957–1978, 2004. doi: https://doi.org/10.1111/j.1540-6261.2004.00688.x.

K. Judd. *Numerical Methods in Economics*. Scientific and Engineering. MIT Press, 1998. ISBN 9780262100717.

K. Judd and S.-M. Guu. Perturbation Solution Methods for Economic Growth Models. In H. Varian, editor, *Economic and Financial Modeling with Mathematica*, pages 80–103. Springer Verlag, 1993. ISBN 9783540978824.

K. Judd and S.-M. Guu. Asymptotic methods for asset market equilibrium analysis. *Economic Theory*, 18(1):127–157, 2001. ISSN 1432-0479.

A. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1913210.

A. Kyle. Informed speculation with imperfect competition. *The Review of Economic Studies*, 56(3):317–355, 1989. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/2297551.

N. Lambert, M. Ostrovsky, and M. Panov. Strategic trading in informationally complex environments. *Econometrica*, 86(4):1119–1157, 2018. doi: 10.3982/ECTA12635.

H. Lan and A. Meyer-Gohde. Solvability of perturbation solutions in DSGE models. *Journal of Economic Dynamics and Control*, 45:366 – 388, 2014. ISSN 0165-1889.

O. Levintal. Fifth-order perturbation solution to DSGE models. *Journal of Economic Dynamics and Control*, 80:1 – 16, 2017. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2017.04.007.

S. Lott. Perturbations in dsge models: An odd derivatives theorem. *Journal of Economic Dynamics and Control*, 106:103722, 2019. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2019.103722. URL https://www.sciencedirect.com/science/article/pii/S0165188919301228.

L. Menkhoff. The use of technical analysis by fund managers: International evidence. *Journal of Banking & Finance*, 34(11):2573–2586, 2010. ISSN 0378-4266. doi: https://doi.org/10.1016/j.jbankfin.2010.04.014.

E. Miller. Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, 32(4): 1151–1168, 1977. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2326520.

S. Moeller, F. Schlingemann, and R. Stulz. How Do Diversity of Opinion and Information Asymmetry Affect Acquirer Returns? *The Review of Financial Studies*, 20(6):2047–2078, 2007. ISSN 0893-9454. doi: 10.1093/rfs/hhm040.

C.-H. Park and S. Irwin. What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, 21(4):786–826, 2007. doi: https://doi.org/10.1111/j.1467-6419.2007.00519.x.

J. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964. ISSN 00129682, 14680262. doi: http://www.jstor.org/stable/1913738.

M. Rostek and M. Weretka. Price inference in small markets. *Econometrica*, 80(2):687–711, 2012. doi: 10.3982/ECTA9573.

G. Rudebusch and E. Swanson. Examining the bond premium puzzle with a dsge model. *Journal of Monetary Economics*, 55:S111 – S126, 2008. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2008.07.007.

R. Sadka and A. Scherbina. Analyst disagreement, mispricing, and liquidity*. *The Journal of*

Finance, 62(5):2367–2403, 2007. doi: https://doi.org/10.1111/j.1540-6261.2007.01278.x.

S. Schmitt-Grohe and M. Uribe. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, 28(4):755 – 775, 2004. ISSN 0165-1889.

E. Swanson, G. Anderson, and A. Levin. Higher-order perturbation solutions to dynamic, discrete-time rational expectations models. *Federal Reserve Bank of San Francisco Working Paper Series*, 01, 2006.