The Ronald O. Perelman Center for
Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
18-011

Perturbations in DSGE Models: Odd Derivatives Theorem

SHERWIN LOTT
University of Pennsylvania
Department of Economics

May 21, 2018

https://ssrn.com/abstract=3190497

Perturbations in DSGE Models: Odd Derivatives Theorem*

Sherwin Lott[†]

May 21, 2018

Abstract

This paper proves a generalization of previous results in the perturbation literature. Perturbation methods compute policy functions to DSGE models using a multivariate Taylor series with respect to the state variables $\mathbf{x}$ and a *perturbation parameter* σ. Schmitt-Grohé and Uribe (2004) shows that Taylor coefficients of order $\mathbf{x}^0\sigma^1$ and $\mathbf{x}^1\sigma^1$ are zero. Andreasen (2012) extends this to order $\mathbf{x}^2\sigma^1$, and shows the $\mathbf{x}^0\sigma^3$ coefficient is zero if innovations are symmetric.

We show that Taylor coefficients of order $\mathbf{x}^r\sigma^1$ are zero for all r. Most generally, if odd moments of the innovations are zero up to some moment $\bar{s}$, then coefficients of order $\mathbf{x}^r\sigma^s$ are zero for all r and odd $s \leq \bar{s}$. (The intuition for this comes from classical portfolio theory.) Eliminating these coefficients significantly reduces what needs to be computed and thereby runtime, memory usage, and numerical errors.

Keywords: Perturbation methods, DSGE models, odd derivatives, computational macroeconomics.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are a primary workhorse of macroeconomics. Since DSGE models are generally intractable, a variety of approaches

*I thank Jesús Fernández–Villaverde for posing this theorem, as well as mentoring myself through the literature and writing process.

†University of Pennsylvania Department of Economics, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA. Email: lotts@sas.upenn.edu. Declarations of interest: none.

have been developed to compute their solutions. One popular way is with *perturbation methods*. A perturbation approximates the solution by a multivariate Taylor series with respect to the state variables $\mathbf{x}$ and a *perturbation parameter* σ. The purpose of this paper is to prove a theorem that many of these Taylor coefficients are zero, which will increase the speed and accuracy of perturbations.

These methods have been well established in other fields, such as the natural sciences, but only relatively recently gained prevalence within economics. Judd and Guu (1993) show how perturbations can be effectively applied to basic growth models. Later work—Judd (1998), Judd and Guu (2001), and Jin and Judd (2002)—build up more of the theoretical foundations.

Perturbations have become standard because they are quite fast, yet maintain reasonable accuracy. (Caldara et al. (2012) run detailed comparisons against Chebyshev polynomials and value function iterations in a real business cycle model with Epstein–Zin preferences and stochastic volatility.) For applications that perturbations already work well on, our computational advances may only be of marginal importance.

A growing number of problems require high order perturbations. Fernández-Villaverde et al. (2011) shows 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. Unfortunately, computation becomes successively much more difficult with each order: errors grow rapidly in magnitude, and runtime can be a binding constraint.

Swanson et al. (2006) concludes that numerical accuracy is especially important for high order perturbations. This is because numerical errors propagate relatively quickly as higher order perturbations build on the solutions to past orders. In their example, "coefficient errors as large as 10% become quite common by about the fourth or fifth order." Numerical errors arise when computing derivatives to the equilibrium conditions. For instance, a model with Epstein-Zin utility can have large non–integer exponents that are invariably computed with numerical error. This then leads to errors in the computed coefficients.

Our theorem shows that many coefficients can be set exactly to zero, which eliminates all errors that come from computing these coefficients. Further, this allows the remaining coefficients to be computed with greater accuracy. The eliminated errors will no longer enter into the remaining coefficients through the equilibrium conditions.

Relatedly, there are fewer coefficients for errors to propagate through. The causes of inaccuracy in high order perturbations are the reasons why reducing the number of coefficients is valuable.

Because runtimes can be substantial, there is now greater focus on computing high order perturbations faster. Levintal (2017) recently developed a "new notation" that sped computation up by manyfold when implemented for fifth–order perturbations. Our theorem fits into the literature through its direct improvement to computation by reducing the number of coefficients that need to be solved for. Further, it could complement Levintal's type of approach. Eliminating coefficients substantially simplifies the equilibrium conditions that determine the remaining coefficients, which new approaches could take advantage of.

Finally, our result is of theoretical interest in understanding the odd σ–order coefficients that it eliminates. Interpreting these coefficients within specific applications would be worthwhile—as has been done with the previous results that we are generalizing. However, we feel doing so in this paper would distract from the broad conclusion of our theorem.

1.1 Odd σ–Order Derivatives

Schmitt-Grohé and Uribe (2004) prove that Taylor coefficients of order $\mathbf{x}^0\sigma^1$ and $\mathbf{x}^1\sigma^1$ are zero (assuming without loss of generality that innovations have mean zero). Andreasen (2012) extends this result to order $\mathbf{x}^2\sigma^1$, and shows the $\mathbf{x}^0\sigma^3$ coefficient is zero if the third moment of the innovations is zero.[1]

The purpose of our paper is to generalize these previous results. We show that if innovations have mean zero, then Taylor coefficients of order $\mathbf{x}^r\sigma^1$ are zero for all r. Most generally, if odd moments of the innovations are zero up to some moment $\bar{s}$, then coefficients of order $\mathbf{x}^r\sigma^s$ are zero for all r and odd $s \leq \bar{s}$.

An implication of this result is an open conjecture stated in Fernández-Villaverde et al. (2016) (page 559). If the innovations are symmetric, then all Taylor coefficients involving an odd σ–order are zero. Because symmetric innovations are widely used, this may be the most applicable implication of our theorem. Note that symmetry is not as strong of an assumption as it may seem since there is little restriction on how state variables enter into the DSGE model. However, models for tail events such as

[1] The statement of his theorem assumes symmetry. However, Andreasen notes in the proof that he just uses that the third moment of the innovations are zero, which is an implication of symmetry.

catastrophes will generally not be symmetric.

Intuition for this proof derives from a classical portfolio theory result. Agents are marginally risk neutral when they have no risk in their portfolio. That is a statement about the first derivative of utility with respect to risk. However, the result more generally applies to all odd order derivatives when the corresponding odd moment of the innovations is zero. This ties back into the DSGE model through the value function. If the value function is unaffected by odd order changes in risk, then the optimal decision rule should not be affected either.

The rest of the paper is organized as follows. In section 2, we detail the standard perturbation setup in the macroeconomics literature. Section 3 develops new arguments and proves the theorem. Section 4 analyzes the computational impact of this, and section 5 concludes.

2 Perturbation Setup

The idea of a perturbation is to approximate solutions over the space of models, which then lets us back out the solution to a particular model. This is conceptually much like a Taylor series, but where the domain is over models. We start with a simple model, in that it has a known or easily computable solution, and then build an approximation around it.

In our macroeconomic context, we want to compute the equilibrium policy functions of a specific DSGE model. To that end, consider the space over all transformations of this model where the innovations 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, centered at this steady–state, with respect to the *state variables* $\mathbf{x}$ and the perturbation parameter σ. The equilibrium to our unscaled model, $\sigma = 1$, can then be backed out. This will be detailed in the rest of the section.

2.1 Standard Setup

We will 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

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

the equilibrium conditions be expressible as a system of equations:

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

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

$$\mathbf{y}_t = \mathbf{g}\left(\mathbf{x}_t; \sigma\right), \tag{2}$$

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t; \sigma\right) + \sigma \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+1} \tag{3}$$

The state variables evolve stochastically. Here, $\boldsymbol{\epsilon}_{t+1}$ is a vector of n_ϵ independent innovations; $\boldsymbol{\eta}$ is a $n_x \times n_\epsilon$ matrix—it linearly transforms the innovations into state variable shocks. Notice 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 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:

$$\mathcal{H}\left(\bar{\mathbf{y}}, \bar{\mathbf{y}}, \bar{\mathbf{x}}, \bar{\mathbf{x}}\right) = 0 \tag{4}$$

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

Here, the purpose of perturbation methods is to find the policy functions of the unscaled model—when $\sigma = 1$. These are $\mathbf{g}\left(\,\cdot\,, 1\right)$ and $\mathbf{h}\left(\,\cdot\,, 1\right)$. They can be backed out 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 (2) and (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 a 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. The unknowns are the derivatives of $\mathbf{g}$ and $\mathbf{h}$ up to the n^{th}–order. This system has exactly as many equations as unknowns—allowing us to compute the Taylor expansion.

2.2 Stochastic Volatility Setup

We can add stochastic volatility into the model by changing how state variables evolve in (3) to the following.

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t;\sigma\right) + \sigma\boldsymbol{\eta}\boldsymbol{\alpha}_{t+1}\boldsymbol{\epsilon}_{t+1} \tag{3*}$$

Here, $\boldsymbol{\alpha}_{t+1}$ is the stochastic volatility term—an $n_\epsilon \times n_\epsilon$ matrix.

All our results will hold for stochastic volatility so long as expectations in the system of equations are still well defined. The stochastic volatility term $\boldsymbol{\alpha}$ may depend on the past $\boldsymbol{\epsilon}$'s in any way. For example, $\boldsymbol{\alpha}$ can be used to represent any GARCH model. The standard setup in section 2.1 corresponds to $\boldsymbol{\alpha}$ being the identity matrix.

Just like $\boldsymbol{\eta}$, the $\boldsymbol{\alpha}_{t+1}$ is a constant with respect to the expectation after time t, and it does not depend on σ. The proofs will go through by simply replacing every instance of $\boldsymbol{\eta}$ with $\boldsymbol{\eta}\boldsymbol{\alpha}_{t+1}$. Therefore, we stick with the notation in section 2.1 by omitting $\boldsymbol{\alpha}_{t+1}$.

3 Main Developments

This section proves the main theorem that says when odd σ–order coefficients are zero. (In the following, functions are evaluated at the deterministic steady state, $(\bar{\boldsymbol{x}};0)$, unless otherwise noted. Let $\boldsymbol{\epsilon}^s$ denote the Kronecker product of $\boldsymbol{\epsilon}$ with itself s times.)

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

This immediately generalizes the results in Schmitt-Grohé and Uribe (2004) and Andreasen (2012), as well as proves the open conjecture in Fernández-Villaverde et al. (2016).

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

Corollary 2 *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*.*

Since this paper is about proving a specific property of perturbations, we will take as given that the standard methods are valid. (See Lan and Meyer-Gohde (2014) for solvability conditions.) That means we assume the system of equations has a unique solution. For this system of equations to even exist, we must assume the expectations are well defined, and that the functions $\mathcal{H}$, $\mathbf{g}$, and $\mathbf{h}$ are differentiable up to whatever order the perturbations are taken to. If any of these assumptions are violated in an application, then perturbations should not be used in the first place.

The intuition for the theorem is straightforward. A classical portfolio theory result says that if $\mathbb{E}[\boldsymbol{\epsilon}^s] = 0$ for all odd $s \leq \bar{s}$, then $\frac{\partial^s}{\partial \sigma^s} V(\mathbf{x}; 0) = 0$ for all $\mathbf{x}$ and odd $s \leq \bar{s}$. (Where V is a representative agent's value function). It then stands to reason that if an agent's utility is marginally unaffected by these odd high order changes in risk, then perhaps their optimal decision is not affected either. That is, $\mathbf{g}_{\sigma^s}(\boldsymbol{x}; 0) = \mathbf{h}_{\sigma^s}(\boldsymbol{x}; 0) = 0$ for all $\mathbf{x}$ and odd $s \leq \bar{s}$. This would then imply the theorem.

While our proof only uses basic calculus, it is a bit more convoluted than the intuition because policy functions are determined by a system of equations (outlined in section 2). The following subsection 3.1 carefully examines the system of equations and proves a basic "bookkeeping" result. Finally, subsection 3.2 uses this to prove the theorem by induction.

3.1 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:[3]

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

Any order derivative can readily be 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 $\boldsymbol{\eta\epsilon}$.[4] *The total number of products being summed over is $J^{(r,s)}$, with j indexing over these products.*

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

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

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

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

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

(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 these terms changed inline 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* $P^{(r,s,j)}$*,* $1 \leq i \leq K^{(r,s,j)}$*. And, denote the sum by,* $b^{(r,s,j)} = \sum\limits_{i=1}^{K^{(r,s,j)}} k_i^{(r,s,j)}$*.*

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

Similarily, 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.

3.2 Proof of Theorem

With these bookkeeping results in mind, we now turn to partially solving the system of equations generated by $F_{\mathbf{x}^r \sigma^s} = 0$. 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 the expectation. 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.[5] All other terms in $F_{\mathbf{x}^r\sigma^s}$ 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. The theorem can now be proven by induction.

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

Proof. We'll prove this by induction on the order of the perturbation. Schmitt-Grohé and Uribe (2004) already proved that $\mathbf{g}_\sigma = \mathbf{h}_\sigma = 0$, which is our base case.

Suppose this theorem holds for policy derivatives in a n^{th}–order perturbation; we want to show that then it holds for policy derivatives in a $(n+1)^{th}$–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 of equations with a unique solution; therefore, it is in fact a partial solution. ■

As noted in the beginning of section 3, we assume there is a unique solution to each perturbation. Otherwise, these perturbation methods should not be used in the first place. While this is an important foundational question, it is outside the purview of this paper, so we consider it a rather innocuous assumption for our purposes. (See Lan and Meyer-Gohde (2014) for solvability conditions.)

[5] It does not matter the order the $\boldsymbol{\epsilon}$'s appear. Every element in the resultant tensor will contain that many elements of $\boldsymbol{\epsilon}$, which will evaluate to zero in expectation if there is an odd such number $\leq \bar{s}$.

4 Analysis

High order perturbations with many state variables require a lot of computing power and memory. This is because the number of coefficients in a perturbation grows *exponentially* with its order n and number of state variables n_x. (See Appendix A.1 for a closed form expression and analysis of the number of coefficients in a perturbation.) Runtime and memory become binding constraints when there are so many coefficients—the computer simply runs out of memory. Further, numerical errors compound as high order perturbations build on lower order solutions (Swanson et al. (2006)).

The primary purpose of our paper is to improve computation by reducing the number of coefficients. Theorem 1 proved that odd σ–order coefficients are zero when the corresponding odd moments of $\boldsymbol{\epsilon}$ are zero. (If $\boldsymbol{\epsilon}$ is symmetric, then all odd σ–order coefficients are zero.) And, Lemma 2 proved that this partial solution solves all the equations of corresponding odd σ–order. We can now eliminate a sizable percentage of the coefficients and equations in perturbations (see Table 1).

These results reduce runtime, memory use, and numerical errors in the computation of perturbations.[6] This is because the coefficients and equations we eliminate no longer need to be computed or stored. Further, these coefficients are now known with perfect accuracy—they are exactly zero. The magnitude of these computational benefits correspond with the percent of coefficients that are eliminated, which is substantial.

In addition, it is now easier to compute the coefficients and equations that remain. The equations are sums of products, and any product that contains an eliminated coefficient is itself zero.[7] Setting these products to zero greatly simplifies the equations, which allows the coefficients to be computed faster and with fewer numerical errors.

A natural followup question is: what proportion of coefficients and equations are eliminated? That is, what proportion of coefficients are of "odd σ–order?" We quantify this in section 4.1. Then, section 4.2 discusses how to modify current methods to best implement our results.

[6]These results will ideally be implemented at a developer level for softwares such as Dynare. Section 4.2 discusses at a high-level how to code this, though we have not done so. We want to keep our general result separate from important coding and application specific details of how best to compute a perturbation. Fruitful future work includes not just implementing our result, but figuring out additional implications. (For instance, we can eliminate many products in the even σ–order equations. How can this fact be utilized? How does it affect the interpretation of various coefficients?)

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

4.1 Proportion of Coefficients that are of Odd σ–Order

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.
$\Big($These are approximately $\frac{n}{2n+n_x}$, see Claim 3 in Appendix A.$\Big)$

See Appendix A for the mathematical details of this section. 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: $\boldsymbol{g}_{x_1}$, $\boldsymbol{g}_{x_2}$, $\boldsymbol{g}_{x_1x_2}$, $\boldsymbol{g}_{x_1^2}$, $\boldsymbol{g}_{x_2^2}$, $\boldsymbol{g}_\sigma$, $\boldsymbol{g}_{x_1\sigma}$, $\boldsymbol{g}_{x_2\sigma}$, 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 σ^3... 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.

4.2 Improving Perturbation Methods

How can our results be used to improve perturbation methods? First, the coefficients and equations eliminated no longer need to be computed. Secondly, the remaining equations (other than those of order σ^0) can be greatly simplified. This reduces computation and improves accuracy.

We will now explain in detail a modified version of the standard *block recursion* method (Jin and Judd (2002)). For a fixed s, *block–s* is the set of equations of order σ^s. In Figure 1, block–s is depicted by the entire row s. The blocks are computed recursively starting with block–0.

Block–0 is computed in the following way. The system of equations $F_{\mathbf{x}^1\sigma^0}$ is computed by differentiating $F_{\mathbf{x}^0\sigma^0}$ and then solved to get the coefficients of order $\mathbf{x}^1\sigma^0$. Similarly, $F_{\mathbf{x}^2\sigma^0}$ is computed by differentiating $F_{\mathbf{x}^1\sigma^0}$ and then solved to get the coefficients of order $\mathbf{x}^2\sigma^0$. The rest of the block can be computed in this way.

By Theorem 1, block–1 can be skipped since all coefficients of order σ^1 are zero. We can then proceed to compute block–2. The system of equations $F_{\mathbf{x}^0\sigma^2}$ is computed by twice differentiating $F_{\mathbf{x}^0\sigma^0}$.

Any product in $F_{\mathbf{x}^0\sigma^2}$ that contains either a coefficient of odd σ–order or an odd number of $\boldsymbol{\epsilon}$'s is itself zero. Further, the derivative of such a product with respect to $\mathbf{x}$ will still be zero because the σ–orders and number of $\boldsymbol{\epsilon}$'s are unchanged (see (D5)–(D8)). Hence, we can eliminate all such products in the computation of block–2.

$$
\begin{tikzcd}
0 & F_{\mathbf{x}^0\sigma^0} \arrow[r] \arrow[d] & F_{\mathbf{x}^1\sigma^0} \arrow[r] & F_{\mathbf{x}^2\sigma^0} \arrow[r] & \cdots \arrow[r] & F_{\mathbf{x}^{n-2}\sigma^0} \arrow[r] & F_{\mathbf{x}^{n-1}\sigma^0} \arrow[r] & F_{\mathbf{x}^{n}\sigma^0} \\
1 & F_{\mathbf{x}^0\sigma^1} \arrow[d] & F_{\mathbf{x}^1\sigma^1} & F_{\mathbf{x}^2\sigma^1} & \cdots & F_{\mathbf{x}^{n-2}\sigma^1} & F_{\mathbf{x}^{n-1}\sigma^1} & \\
2 & F_{\mathbf{x}^0\sigma^2} \arrow[r] \arrow[d] & \hat{F}_{\mathbf{x}^1\sigma^2} \arrow[r] & \hat{F}_{\mathbf{x}^2\sigma^2} \arrow[r] & \cdots \arrow[r] & \hat{F}_{\mathbf{x}^{n-2}\sigma^2} & & \\
3 & F_{\mathbf{x}^0\sigma^3} \arrow[d] & F_{\mathbf{x}^1\sigma^3} & F_{\mathbf{x}^2\sigma^3} & \cdots & & & \\
 & \vdots \arrow[d] & & \cdots & & & & \\
n & F_{\mathbf{x}^0\sigma^n} & & & & & &
\end{tikzcd}
$$

Figure 1: Modified block recursion.

This substantially reduces the number of products that need to be kept track of and differentiated. Denote this simplified system of equations by $\hat{F}_{\mathbf{x}^0\sigma^2}$.

Block–2 can now be computed in the same way as block–0, just with $\hat{F}_{\mathbf{x}^0\sigma^2}$. That is, $\hat{F}_{\mathbf{x}^0\sigma^2}$ is solved to get the coefficients of order $\mathbf{x}^0\sigma^2$. Then, $\hat{F}_{\mathbf{x}^1\sigma^2}$ is computed by differentiating $\hat{F}_{\mathbf{x}^0\sigma^2}$ and then solved to get the coefficients of order $\mathbf{x}^1\sigma^2$, and so on.

The blocks can continue to be solved in this recursive way. The odd blocks are skipped so long as the corresponding odd moments of $\boldsymbol{\epsilon}$ are zero. (All of the odd blocks can be skipped if $\boldsymbol{\epsilon}$ is symmetric.) The equations in the remaining (even) blocks are computed normally—except that the simplified system of equations $\hat{F}_{\mathbf{x}^0\sigma^s}$ is used.

We have not changed how the deterministic block–0 is computed. What we have done is eliminated odd blocks and greatly simplified all the blocks that remain besides block–0.

5 Conclusion

We proved in Theorem 1 that if odd moments of innovations are zero up to some moment $\bar{s}$, then coefficients of order $\mathbf{x}^r\sigma^s$ are zero for all r and odd $s \leq \bar{s}$. This is a generalization of the theoretical results in Schmitt-Grohé and Uribe (2004) and Andreasen (2012) to all orders. In doing so, we also proved an open conjecture in Fernández-Villaverde et al. (2016) that all coefficients of an odd σ–order are zero when the innovations are symmetric.

The proportion of coefficients and equations eliminated (when $\boldsymbol{\epsilon}$ is symmetric) is given by Table 1 in section 4.1. Since this portion of the perturbation no longer needs to be computed or stored, we expect reduction in runtime, memory usage, and numerical errors to be comparable. This is significant for high order perturbations where computing power and memory can be binding constraints, and accuracy issues are acute (Swanson et al. (2006)).

In addition, the remaining equations can be greatly simplified. This is because any product containing an eliminated coefficient is itself zero. Setting these products to zero simplifies the equations, which further allows the coefficients to be computed faster, with less memory, and fewer numerical errors.

Beyond computational improvements, Theorem 1 enhances our understanding of perturbations. We now know that odd σ–order coefficients are zero and how it relates to classical portfolio theory.

References

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

Caldara, Dario, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao (2012), "Computing DSGE models with recursive preferences and stochastic volatility." *Review of Economic Dynamics*, 15, 188 – 206.

de Groot, Oliver (2015), "Solving asset pricing models with stochastic volatility." *Journal of Economic Dynamics and Control*, 52, 308 – 321.

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

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martin Uribe (2011), "Risk Matters: The Real Effects of Volatility Shocks." *The American Economic Review*, 101, 2530 – 2561.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, and Frank Schorfheide (2016), "Chapter 9 - Solution and Estimation Methods for DSGE Models." In *Handbook of Macroeconomics* (John B. Taylor and Harald Uhlig, eds.), volume 2, 527 – 724, Elsevier.

Jin, He-Hui and Kenneth L. Judd (2002), "Perturbation methods for general dynamic stochastic models." *Manuscript, Stanford University.*

Judd, Kenneth L. (1998), *Numerical Methods in Economics*. Scientific and Engineering, MIT Press.

Judd, Kenneth L. and Sy-Ming Guu (1993), "Perturbation Solution Methods for Economic Growth Models." In *Economic and Financial Modeling with Mathematica* (H.R. Varian, ed.), 80–103, Springer Verlag.

Judd, Kenneth L. and Sy-Ming Guu (2001), "Asymptotic methods for asset market equilibrium analysis." *Economic Theory*, 18, 127–157.

Lan, Hong and Alexander Meyer-Gohde (2014), "Solvability of perturbation solutions in DSGE models." *Journal of Economic Dynamics and Control*, 45, 366 – 388.

Levintal, Oren (2017), “Fifth-order perturbation solution to DSGE models.” *Journal of Economic Dynamics and Control*, 80, 1 – 16.

Schmitt-Grohé, Stephanie and Martín Uribe (2004), “Solving dynamic general equilibrium models using a second-order approximation to the policy function.” *Journal of Economic Dynamics and Control*, 28, 755 – 775.

Swanson, Eric, Gary Anderson, and Andrew Levin (2006), “Higher-order perturbation solutions to dynamic, discrete-time rational expectations models.” *Federal Reserve Bank of San Francisco Working Paper Series*, 01.

A Computational Analysis

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.

A.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 2 gives an exact closed form expression for the number of coefficients in such an order n perturbation.

Claim 2 *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).

A.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.

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

Here, we are summing over the number of coefficients with σ–order s. For a fixed s, the number of nonnegative integer solutions to $r_1 + r_2 + \ldots + r_{n_x} \le n - s$ 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.

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

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 3 *The proportion of coefficients of odd σ–order is bounded between:*

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

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

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,

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

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.

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

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

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

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 $\hat{p}_{odd}(n, n_x)$.

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

■