The Use of Symmetry for Models with Variable-size Variables

Takeshi Fukasawa*

May 2, 2024

Abstract

This paper shows the universal representations of symmetric functions with multidimensional variable-size variables, which help assessing the justification of approximation methods aggregating the information of each variable by moments. It then discusses how the results give insights into economic applications, including two-step policy function estimation, moment-based Markov equilibrium, and aggregative games.

Keywords: Symmetry; Permutation invariance; Functions with variable-size variables; Approximation

*Graduate School of Public Policy, The University of Tokyo. 7-3-1 Hongo, Tokyo, Japan. E-mail: fukasawa3431@gmail.com.
This paper is based on Chapter 6 of my Ph.D. dissertation at the University of Tokyo. I thank Hiroshi Ohashi, Kei Kawai, and my dissertation committiee members for their comments.
This work was supported by JSPS KAKENHI Grant Number JP21J21410.

1 Introduction

Many economic problems require approximation methods, because of their complexity and computational burden. Examples include two-step policy function estimation methods for dynamic models (Hotz and Miller, 1993, Bajari et al., 2007),[1] and Moment-based Markov Equilibrium (MME; Ifrach and Weintraub, 2017) as an approximation of Markov perfect Equilibrium (MPE). In these methods, precise approximations of functions (e.g., value functions, policy functions) are essential for deriving correct implications.

One obstacle to the approximation of functions is the existence of variable-size variables. Consider a dynamic investment competition model as in Ericson and Pakes (1995), where all the firms' capital stocks at the beginning of each period are the state variables. Policy function corresponds to the density of the value of each firm's investment as a function of states. Then, how should we estimate the policy function using data in multiple markets in reduced form, if there are three firms in some markets, but two in other markets? Here, the states are variable-size, because the size of states is heterogeneous across markets.

This study shows the universal representations of multidimensional symmetric functions with variable-size variables using polynomial functions (Section 3), building on the machine learning and mathematics literature since Zaheer et al. (2017). The results help assessing the justifications of approximation methods aggregating the information of each variable by moments. It then discusses how the results give insights into economic applications, including two-step policy function estimation, MME, and aggregative games[2] (Section 4). Regarding the example of the policy function estimation discussed above, it is justifiable to estimate a common policy function as a function of own firm's states and the sums of polynomial terms (moments) of competitors' states under some conditions, regardless of the number of firms in each market, as long as the number of moments is sufficiently large. Concerning the MME, this study shows that MME is equivalent to the MPE if the number of moments reaches a certain level. Regarding aggregative games, we can easily show that any games can be represented as multidimensional generalized aggregative, which introduces multidimensional aggregates in the generalized (fully) aggregative games (Cornes and Hartley, 2012).

The rest of this paper is organized as follows. Section 2 describes the relations to the previous studies. Section 3 shows the mathematical results, and Section 4 discusses the economic application. Section 5 concludes.

Appendix A shows all the proofs of the main propositions. In Appendix B, we further discuss the implications for models with dynamic demand and multi-product firms.

2 Literature

First, this study relates to and contributes to economic studies based on game theoretic models. It is sometimes not easy to directly estimate or solve the models, and some approximation techniques in a broad sense are introduced (e.g., two-step policy function estimation method, MME). By showing general mathematical results based on the notion of symmetry, the current study evaluates the justification of these methods from a different point of view. The mathematical results are very general and simple, and they would be useful for further economic research.

Note that the idea of using symmetry is not new in the literature. For instance, Pakes and McGuire (1994) discussed using symmetry to reduce the dimension of state variables in a dynamic game with finite state space. Nevertheless, the strategy is applicable only to the model with fixed-size finite state space. In contrast, the mathematical results in the current paper can be also applied to dynamic models with both fixed-size/variable-size and discrete/continuous state space, and they are more general.

Kahou et al. (2021) also discussed using symmetric structures with fixed-size variables, mainly for quantitatively solving heterogeneous agent macroeconomic models using deep learning techniques. The current study complements their analysis by formally showing the representations allowing for the case of variable-size variables. Note that they speculated in Section 6.3 that exploiting the symmetric structure enables researchers to solve dynamic models with networks. Typically, network structure is characterized by variable-size variables, because the number of each agent's neighbors is heterogeneous.

[1] In two-step estimation methods, policy functions used in the second step are "approximated" by the ones consistently estimated in reduced-form in the first step.

[2] The latter two consider the settings with fixed-size variables.

Finally, this study builds on the machine learning and mathematics literature on symmetric functions (e.g., Zaheer et al., 2017). As discussed in Section 3, Wagstaff et al. (2022) also derived a representation of symmetric functions with single-dimensional variable-size variables. Nevertheless, the construction is not intuitive, and not easy to interpret in applications, especially economics. The current study shows a more intuitive representation of symmetric functions with variable-size variables using moments.

3 Universal representation of symmetric functions with variable-size variables

First, we define permutation invariance, which is one kind of symmetry.

Definition 1. A function $V : \Omega^{J'} \subset \mathbb{R}^{I\times J'} \to \mathbb{R}$ is permutation invariant,[3] if, for all permutations $\pi \in \mathcal{S}_{J'}$, $V(\pi(x_1,\cdots,x_{J'})) = V(x_1,\cdots,x_{J'})$. A function $V : \cup_{J'\leq J}\Omega^{J'} \subset \mathbb{R}^{I\times(\leq J)} \to \mathbb{R}$ is permutation invariant if $V|_{\Omega^{J'}}$ is permutation invariant for every $J' \leq J$.

For instance, a fixed-size function $f : \mathbb{R}^3 \to \mathbb{R}$ such that $f(x_1,x_2,x_3) = x_1x_2+x_1x_3+x_2x_3$ and a variable-size function $g : (\mathbb{R}^2 \cup \mathbb{R}^3) \to \mathbb{R}$ such that $g(x_1,x_2) = x_1x_2$, $g(x_1,x_2,x_3) = x_1x_2+x_1x_3+x_2x_3$ are permutation invariant. In the rest of this paper, let $V\left(\{x_j\}_{j\in\mathcal{J}}, y\right)$ be the permutation invariant function whose output is invariant to the order of the elements $x_j(j\in\mathcal{J})$, but not necessarily on the order of $\{x_j\}_{j\in\mathcal{J}}$ and y. Besides, let $(x_1,\cdots,x_J)$ be a tuple with J elements. In the tuple, the order of elements matters.

The following definition is on the continuity of functions with variable-size variables:

Definition 2. A function $V : \cup_{J'\leq J}\Omega^{J'} \subset \mathbb{R}^{I\times(\leq J)} \to \mathbb{R}$ is continuous if $V|_{\Omega^{J'}}$ is continuous for every $J' \leq J$.

Next, we define the function called (multi)symmetric power sum:

Definition 3. $\eta_{I,J} : [0,\infty)^I \to \mathbb{R}^{\kappa(I,J)}$ $\left(\kappa(I,J) \equiv \begin{pmatrix} J+I \\ I \end{pmatrix} - 1\right)$, defined by

$$\eta^{(s)}_{I,J}(x) = x^{s_1}_{i=1}x^{s_2}_{i=2}\cdots x^{s_I}_{i=I} \quad s \in \left\{(s_1,s_2,\cdots,s_I) \in \mathbb{Z}^I_+ | 1 \leq s_1+s_2+\cdots+s_I \leq J\right\}$$

is called multisymmetric power sum[4] of degree up to J in I variables.

In the special case $I = 1$, it is equivalent to polynomials up to J-th order: $x, x^2, \cdots, x^J$.

In the following, let $\min\Omega \equiv \{x \in \Omega | x \leq x' \ \forall x' \in \Omega\}$ and $\Omega^{\leq J} \equiv \cup_{J'\leq J}\Omega^{J'}$. Then, we obtain the following statement:[5]

Proposition 1. *Let Ω be a compact subset of $[0,\infty)^I$. Suppose either of the following conditions holds for a unique continuous permutation invariant function $V : \Omega^{\leq J} \to \mathbb{R}$:*

(a). $\min\Omega > 0_I$

(b). $V\left(x_1,\cdots,x_{J'},\underbrace{0,\cdots,0}_{J-J'}\right) = V(x_1,\cdots,x_{J'})$

Then, there exists a continuous function $\psi : \mathbb{R}^{\kappa(I,J)} \to \mathbb{R}$ such that:

$$V(\{x_j\}_{j\in\mathcal{J}}) = \psi\left(\sum_{k\in\mathcal{J}} \eta_{I,J}(x_k)\right) \ |\mathcal{J}| \leq J.$$

[3] In some literature in industrial organization (e.g., Doraszelski and Pakes, 2007), the property is called anonymity or exchangeability. They discuss the use of symmetry in the dynamic model with discrete states. Our discussion is more general in that we allow for the case with continuous states.

[4] In this definition, we exclude the term such that $s_i = 0 \ \forall i$.

[5] It is trivial that functions in the form of $\psi\left(\sum_{k\in\mathcal{J}}\eta_{I,J}(x_k)\right)$ are permutation invariant. In contrast, whether permutation invariant functions can be represented as $\psi\left(\sum_{k\in\mathcal{J}}\eta_{I,J}(x_k)\right)$ is not necessarily trivial.

Intuitively, condition (b) implies j-th ($j = J'+1, \cdots, J$) variables are negligible in the function V when they take the values of zero, and they have no effect on the values of "aggregated variables" $\sum_{k\in\mathcal{J}} \eta_{I,J}(x_k)$, and consequently the values of V.

Besides, though the range of $\Omega \subset [0,\infty)^I$ might seem to be restrictive,[6] we can transform the domain of functions by applying the appropriate change of variables.[7]

Remark 1. Wagstaff et al. (2022) showed there exist continuous functions $\phi : [0,1] \to Im(\phi) \subset \mathbb{R}^J$ and $\Psi : \left\{\sum_{k\in\mathcal{J}} \phi(x_k); x_k \in [0,1], |\mathcal{J}| \le J\right\} \to \mathbb{R}$ such that $V(\{x_j\}_{j\in\mathcal{J}}) = \Psi\left(\sum_{k\in\mathcal{J}} \phi(x_k)\right)$ $|\mathcal{J}| \le J$ for a continuous permutation invariant function $V : [0,1]^{\le J} \to \mathbb{R}$. Though we can extend the proof to the case with multidimensional variables, their construction of the function ϕ is not intuitive, unlike the case of η in Proposition 1. Our result is easier to understand because we use polynomial functions. Also, by the construction of Ψ and ϕ, the subset of the domain of Ψ, $\left\{\sum_{k\in\mathcal{J}} \phi(x_k); x_k \in [0,1]\right\} \subset \mathbb{R}^J$, should be disjoint for different values of $|\mathcal{J}|$. It implies large domains of Ψ, which would be hard to approximate in practice.

The next result is more relevant to economic applications.

Proposition 2. *Let Ω be a compact subset of $[0,\infty)^I$, and let Υ be a subset of $\mathbb{R}^C$. Suppose either of the following conditions holds for a unique continuous permutation invariant function $V : \Omega \times \Omega^{\le J-1} \times \Upsilon \to \mathbb{R}$:*

(a). $\min \Omega > 0_I$

(b). $V\left(x_j, \left\{x_{j'}\right\}_{j'\in\mathcal{J}-\{j\}}, y\right) = V\left(x_j, \left\{x_{j'}\right\}_{j'\in\mathcal{J}-\{j\}\ s.t.\ x_{j'}\neq 0}, y\right).$

Then, there exists a continuous function $\psi : \Omega \times \mathbb{R}^{\kappa(I,J-1)} \times \Upsilon \to \mathbb{R}$ such that:

$$V\left(x_j, \left\{x_{j'}\right\}_{j'\in\mathcal{J}-\{j\}}, y\right) = \psi\left(x_j, \sum_{k\in\mathcal{J}-\{j\}} \eta_{I,J-1}(x_k), y\right) \quad |\mathcal{J}| \le J.$$

4 Economic Applications

4.1 Two-step policy function estimation

It is known that consistent estimation of policy functions is essential for precise estimation and counterfactual simulation of the empirical models of games, when applying two-step estimation methods (e.g., Bajari et al. (2007)). The propositions give insights into the functional form of policy functions used in estimations.

Here, consider Ericson and Pakes (1995) type dynamic competition model, and suppose firms in all the markets follow the same symmetric Markov perfect equilibrium (MPE), where all the firms in all the markets share the same policy functions $\sigma\left(s_{jm}, \left\{s_{j'm}\right\}_{j'\neq j}, y_m, \nu_{jm}\right)$, and invariant to permutations of competitors' states $\left\{s_{j'm}\right\}_{j'\neq j}$. Here, $s_{jm} \in \Omega \subset [0,\infty)^I$ denotes firm j's states in market m, ν_{jm} denotes the firm's private shock, and y_m denotes market m's market-level states. Let $\mathcal{J}_m$ be the set of firms in market m. Let J be the maximum number of firms in all the markets. Suppose the domain of s_{jm}, Ω, and the policy function, satisfy condition (a) or (b) in Proposition 2. If the firm with $s_{jm} = 0$ has negligible impact on other firms' policy function, condition (b) is satisfied. If not, we can apply change of variables so that condition (a) is satisfied. Then, by the proposition, the policy function is in the following form: $\sigma\left(s_{jm}, \left\{s_{j'm}\right\}_{j'\neq j}, y_m, \nu_{jm}\right) = \exists\widetilde{\sigma}\left(s_{jm}, \sum_{j'\in\mathcal{J}_m-\{j\}} \eta_{I,J-1}\left(s_{j'm}\right), y_m, \nu_{jm}\right)$.

It implies probability or density of choosing a_{jm} at state $\left(s_{jm}, \left\{s_{j'm}\right\}_{j'\neq j}, y_m\right)$ can be represented as $Pr\left(a_{jm}|s_{jm}, \left\{s_{j'm}\right\}_{j'\neq j}, y_m\right) = \exists g\left(s_{jm}, \sum_{j'\in\mathcal{J}_m-\{j\}} \eta_{I,J-1}\left(s_{j'm}\right), y_m\right)$.[8] For instance, in the case $I = 1$, we can represent $Pr\left(a_{jm}|s_{jm}, \left\{s_{j'm}\right\}_{j'\neq j}, y_m\right) = \exists g\left(s_{jm}, \left\{\sum_{j'\in\mathcal{J}_m-\{j\}} \left(s_{j'm}\right)^q\right\}_{q=1,\cdots,J-1}, y_m\right)$, using a common

[6] In the case of fixed variable sizes, Ω can be $\mathbb{R}^I$, and the conditions (a), (b) are not necessary, as shown in Corollary 2.3 of Chen et al. (2022).

[7] For instance, for a state $x \in [-M, M]$ ($M > 0$), we can define an alternative state $\widetilde{x} \equiv \exp(x) \in [\exp(-M), \exp(M)]$.

[8] See Bajari et al. (2007) for the discussion on the continuous choice case.

function g regardless of the number of firms in each market. Hence, we can estimate a policy function as a function of unnormalized moments of competitors' states, even when the number of firms is heterogeneous across markets.[9] Though the number of moments should be $\kappa(I, J)$ in general, we can expect adding higher-order moments yields minor differences. Although it depends on empirical contexts, a small number of moments might be enough to approximate the functions well.

The use of (unnormalized) moments in policy function estimation has been used in the literature.[10] The results formally justify such a strategy, as long as the number of moments is sufficiently large.

4.2 Moment-based Markov equilibrium

Ifrach and Weintraub (2017) proposed Moment-based Markov equilibrium (MME) for dynamic oligopoly models with a small number of dominant firms and a large number of fringe firms, where firms keep track of dominant incumbent firms states and a few moments of fringe incumbent firms' states. It alleviates the computational burden of solving MPE through reduced state space. The results in Section 3 imply symmetric MME is equivalent to MPE when we use a large number of symmetric power sums as moments.[11]

To make the point clear, we omit the existence of dominant firms and firms' entry/exit decisions.[12] Let $x_j \in \mathbb{N}^q$ be firm j's states, such as the capacity or quality level of the firm's product.

Suppose firms follow symmetric MPE. Then, Proposition 2 implies firms' value function can be reformulated as $V\left(x_j, \{x_{j'}\}_{j' \in \mathcal{J}-\{j\}}\right) = \overline{V}\left(x_j, \sum_{j' \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_{j'})\right)$. Similarly, firms' investment strategy can be reformulated as the function of $\sum_{j' \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_{j'})$. Furthermore, Corollary 1 in Appendix A implies there exists an one-to-one mapping between $\{x_j\}_{j \in \mathcal{J}}$ and $\sum_{j' \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_{j'})$. Hence, solving the model with states $\sum_{j' \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_{j'})$ is equivalent to solving the MPE. The former corresponds to solving MME using $\sum_{j' \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_{j'})$ as moments.

When the order of symmetric power sums is less than J, we can expect adding higher order terms provides smaller information in many cases. In such cases, using small number of moments is sufficient to well approximate MPE[13].

4.3 Aggregative games

Section 3 's results also give insights into aggregative games. Let $g_j : \mathcal{A} \to \mathbb{R}$ be player j's payoff function, where $\mathcal{A} \equiv \{\mathcal{A}_j\}_{j \in \mathcal{J}}$, and $\mathcal{A}_j$ denotes player j's strategy set. The following is the definition of aggregative games.

Definition 4. (Cornes and Hartley, 2012) The game is called generalized (fully) aggregative, if there exist functions $\widetilde{g_j} : \mathcal{A}_j \times \mathbb{R} \to \mathbb{R}$ and $h_j : \mathcal{A}_j \to \mathbb{R}$ such that $g_j(a) = \widetilde{g_j}\left(a_j, \sum_{j \in \mathcal{J}-\{j\}} h_{j'}(a_{j'})\right)$.

We can analogously define:

Definition 5. The game is called multidimensional generalized (fully) aggregative, if there exists an integer K and functions $\widetilde{g_j} : \mathcal{A}_j \times \mathbb{R}^K \to \mathbb{R}$ and $h_j : \mathcal{A}_j \to \mathbb{R}^K$ such that $g_j(a) = \widetilde{g_j}\left(a_j, \sum_{j' \in \mathcal{J}-\{j\}} h_{j'}(a_{j'})\right)$.

Then, we obtain the following statement:

Proposition 3. *Suppose player i's payoff function g_j can be written as $g_j(a) = \widetilde{\widetilde{g}}_j\left(a_j, \left\{a_{j'}, x^0_{j'}\right\}_{j' \in \mathcal{J}-\{j\}}\right)$, $\widetilde{\widetilde{g}}_j$ is continuous, and permutation invariant with respect to $\left\{a_{j'}, x^0_{j'}\right\}_{j' \in \mathcal{J}-\{j\}}$. Then, the game is multidimensional generalized aggregative.*

[9]Instead we can separately estimate policy functions for markets with the same number of firms. If the number of observations is small and the number of firms is largely heterogeneous across markets, the strategy might work poorly.

[10]For instance, Ryan (2012) considered a capacity competition model in the cement industry, and estimated firms' investment policy function, as a function of the sum of competitors' capacity.

[11]Ifrach and Weintraub (2017) showed MME becomes an exact approximation of MPE in the constant returns to scale model. Still, the correspondence in more general settings has not been clear.

[12]Essential ideas would not be lost with this simplification.

[13]Ifrach and Weintraub (2017) derived the deviation error bounds when using MME as an approximation of MPE.

Proposition 3 implies all the games satisfying the condition of permutation invariance and continuity of payoff functions are in the form of multidimensional generalized aggregative. Generalized aggregative games are the special cases with $K = 1$.

5 Conclusions

This paper have shown the universal representations of symmetric functions with multidimensional variable-size variables, which helps assessing the justification of approximation methods aggregating the information of each variable by moments. It has also discussed how the results give insights into economic applications, including two-step policy function estimation, moment-based Markov equilibrium, and aggregative games.

Though we have mainly considered three economic applications based on the mathematical results, we would be able to find more applications. I leave it for further research.

References

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Chen, C., Chen, Z., and Lu, J. (2022). Representation theorem for multivariable totally symmetric functions. *arXiv preprint arXiv:2211.15958*.

Cornes, R. and Hartley, R. (2012). Fully aggregative games. *Economics Letters*, 116(3):631–633.

Doraszelski, U. and Pakes, A. (2007). A framework for applied dynamic analysis in IO. *Handbook of industrial organization*, 3:1887–1966.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53–82.

Goettler, R. L. and Gordon, B. R. (2011). Does AMD spur Intel to innovate more? *Journal of Political Economy*, 119(6):1141–1200.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Ifrach, B. and Weintraub, G. Y. (2017). A framework for dynamic oligopoly in concentrated industries. *The Review of Economic Studies*, 84(3):1106–1150.

Kahou, M. E., Fernández-Villaverde, J., Perla, J., and Sood, A. (2021). Exploiting symmetry in high-dimensional dynamic programming. Technical report, National Bureau of Economic Research.

Pakes, A. and McGuire, P. (1994). Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model. *The RAND Journal of Economics*, 25(4):555.

Ryan, S. P. (2012). The Costs of Environmental Regulation in a Concentrated Industry. *Econometrica*, 80(3):1019–1061.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81(5):1763–1803.

Wagstaff, E., Fuchs, F. B., Engelcke, M., Osborne, M. A., and Posner, I. (2022). Universal approximation of functions on sets. *Journal of Machine Learning Research*, 23(151):1–56.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R. R., and Smola, A. J. (2017). Deep sets. *Advances in neural information processing systems*, 30.

A Proof

A.1 Proof of Propositions 1 and 2

Let $\mathcal{M}(I,J;W) \subset M(I,J:\mathbb{R})$ be the set of matrices whose column vector is in $W \subset \mathbb{R}^I$, and the rows are sorted based on the descending lexicographical order.[14] Let a_j be the j-th column vector of the matrix $A \in \mathcal{M}(I,J;W)$, and let a_{ij} be the (i,j)-th element of the matrix A.

Besides, for a compact set $\Omega \subset [0,\infty)^I$, let $\widetilde{\Omega} \subset [0,\infty)^I$ be another compact set such that $\Omega \subset \widetilde{\Omega}$, $\min \widetilde{\Omega} = 0_I$ and $\max \widetilde{\Omega} = \max \Omega$.

Proposition 4. *(Based on Chen et al., 2022)*
Given a compact subset $W \subset \mathbb{R}^I$, there exist a continuous function $\overline{\eta}_{I,J} : W\ (\subset \mathbb{R}^I) \to Im(\overline{\eta})\ (\subset \mathbb{R}^K)$ and a unique homeomorphism $\overline{\Lambda} : \left\{\sum_{j=1}^J \overline{\eta}_{I,J}(a_j) | a_j \in \Omega\right\}\ (\subset \mathbb{R}^K) \to Im(\overline{\Lambda})\ (\subset \mathbb{R}^J)$ such that

$$\overline{\eta}^{(s)}_{I,J}(A_j) = a^{s_1}_{i=1} a^{s_2}_{i=2} \cdots a^{s_I}_{i=I} \quad s \in \{(s_1, s_2, \cdots, s_I) \in \mathbb{Z}_+ | 0 \le s_1 + s_2 + \cdots + s_I \le J\}$$

and

$$A = \overline{\Lambda}\left(\sum_{j=1}^J \overline{\eta}_{I,J}(a_j)\right) \quad \forall A \in \mathcal{M}(I,J;W).$$

The statement can be derived following the proof of Theorem 2.1 of Chen et al. (2022). Using the proposition, we can easily derive the following:

Corollary 1. *Given a compact subset $W \subset \mathbb{R}^I$, there exist a continuous function $\eta_{I,J} : W\ (\subset \mathbb{R}^I) \to Im(\eta)\ (\subset \mathbb{R}^{K-1})$ and a unique homeomorphism $\Lambda : \left\{\sum_{j=1}^J \eta_{I,J}(a_j) | a_j \in \Omega\right\}\ (\subset \mathbb{R}^K) \to Im(\Lambda)\ (\subset \mathbb{R}^J)$ such that*

$$\eta^{(s)}_{I,J}(A_j) = a^{s_1}_{i=1} a^{s_2}_{i=2} \cdots a^{s_I}_{i=I} \quad s \in \{(s_1, s_2, \cdots, s_I) \in \mathbb{Z}_+ | 1 \le s_1 + s_2 + \cdots + s_I \le J\}$$

,and

$$A = \Lambda\left(\sum_{j=1}^J \eta_{I,J}(a_j)\right) \quad \forall A \in \mathcal{M}(I,J;W).$$

Proof. First, we define a homeomorphism $\gamma : \left\{\sum_{j=1}^J \eta_{I,J}(a_j) | a_j \in \Omega\right\} \to \left\{\sum_{j=1}^J \overline{\eta}_{I,J}(a_j) | a_j \in \Omega\right\}$ such that $\gamma(x_1, \cdots, x_{K-1}) = (J, x_1, \cdots, x_{K-1})$. Then, $\sum_{j=1}^J \overline{\eta}_{I,J}(a_j) = \gamma\left(\sum_{j=1}^J \eta_{I,J}(a_j)\right)$ holds because $\sum_{j=1}^J \overline{\eta}^{(s=(0,\cdots,0))}_{I,J}(x_j) = \sum_{j=1}^J 1 = J$, and we have $A = \overline{\Lambda}\left(\sum_{j=1}^J \overline{\eta_{I,J}}(a_j)\right) = \overline{\Lambda}\left(\gamma^{-1}\left(\sum_{j=1}^J \eta_{I,J}(a_j)\right)\right)$. Hence, by defining a function $\Lambda \equiv \overline{\Lambda} \circ \gamma^{-1}$, we obtain the statement. □

The next lemma is used to extend the results above to the case of functions with variable-size variables.

Lemma 1. *For a compact subset $\Omega \subset [0,\infty)^I$ such that $\min \Omega > 0$, let $\widehat{\Omega} \equiv \Omega \cup \prod_{i=1}^I [0, \min \Omega_i]$. For a continuous permutation invariant function $V : \Omega^{\le J} \to \mathbb{R}$, define a function $\overline{V} : \mathcal{M}(I,J,\widehat{\Omega}) \to \mathbb{R}$ such that $\overline{V}(x_1, \cdots, x_J) = \sum_{\{j \in \{1,\cdots,J\}: x_j \in \Omega\} \subset \widetilde{\mathcal{J}} \subset \{1,\cdots,J\}} \left[\prod_{j \in \widetilde{\mathcal{J}}} a(x_j)\right] \cdot \left[\prod_{j \in \{1,\cdots,J\} - \widetilde{\mathcal{J}}} b(x_j)\right] \cdot V\left(\{\max\{x_j, \min \Omega\}\}_{j \in \widetilde{\mathcal{J}}}\right)$,*
where $a(x) \equiv \prod_i \min\left\{1, \frac{x^{(i)}}{\min \Omega^{(i)}}\right\}$ $(x \in [0,\infty)^I)$ and $b(x) \equiv \prod_i \max\left\{0, \frac{\min \Omega^{(i)} - x^{(i)}}{\min \Omega^{(i)}}\right\}$ $(x \in [0,\infty)^I)$.
Then,
(a).$\overline{V}$ is continuous, and
(b).$\overline{V}$ satisfies $\overline{V}\left(x_1, \cdots, x_{J'}, \underbrace{0_I, \cdots, 0_I}_{J-J'}\right) = V\left(\{x_j\}_{j=1,\cdots,J'}\right) \quad (x_j \in \Omega;\ j = 1, \cdots, J').$

[14] The sorting guarantees the uniqueness of $\overline{\Lambda}$ in Proposition 4.

Proof. **Proof of (a).**

It suffices to show $\lim_{x_J \uparrow m} \overline{V}(x_1, \cdots, x_{J-1}, x_J) = \overline{V}(x_1, \cdots, x_{J-1}, x_J = \min \Omega)$.

Because $\lim_{x_J \uparrow m} b(x_J) = b(m) = 0$, the terms associated with $\widetilde{\mathcal{J}}$ such that $J \notin \widetilde{\mathcal{J}}$ disappear when taking the limit:

$$\begin{aligned}
& \lim_{x_J \uparrow m} \overline{V}(x_1, \cdots, x_{J-1}, x_J < m) \\
= & \lim_{x_J \uparrow m} \sum_{\{j \in \{1,\cdots,J-1\}: x_j \in \Omega\} \subset \widetilde{\mathcal{J}} \subset \{1,\cdots,J\}} \left[\prod_{j \in \widetilde{\mathcal{J}}} a(x_j)\right] \cdot \left[\prod_{j \in \{1,\cdots,J\} - \widetilde{\mathcal{J}}} b(x_j)\right] \cdot V\left(\{\max\{x_j, \min \Omega\}\}_{j \in \mathcal{J}}\right) \\
= & \sum_{\{j \in \{1,\cdots,J-1\}: x_j \in \Omega\} \cup \{J\} \subset \widetilde{\mathcal{J}} \subset \{1,\cdots,J\}} \left[\prod_{j \in \widetilde{\mathcal{J}}} a(x_j)\right] \cdot \left[\prod_{j \in \{1,\cdots,J\} - \widetilde{\mathcal{J}}} b(x_j)\right] \cdot V\left(\{\max\{x_j, \min \Omega\}\}_{j \in \mathcal{J}}\right) \\
= & \overline{V}(x_1, \cdots, x_{J-1}, x_J = \min \Omega).
\end{aligned}$$

Hence, we obtain the statement.

Proof of (b).

Let $x_1, \cdots, x_{J'} \in \Omega$. Because $a(0_I) = 0$, the terms associated with $\widetilde{\mathcal{J}} \neq \{j \in \{1, \cdots, J\} : x_j \in \Omega\}$, i.e. $\widetilde{\mathcal{J}} \neq \{1, \cdots, J'\}$ disappear, and

$$\begin{aligned}
& \overline{V}(x_1, \cdots, x_{J'}, 0_I \cdots, 0_I) \\
= & \left[\prod_{j \in \{1,\cdots,J'\}} a(x_j)\right] \cdot \left[\prod_{j \in \{1,\cdots,J\} - \{1,\cdots,J'\}} b(x_j)\right] \cdot V\left(\{\max\{x_j, \min \Omega\}\}_{j \in \mathcal{J}}\right) \\
= & \left[\prod_{j \in \mathcal{J}} 1\right] \cdot \left[\prod_{j \in \{1,\cdots,J\} - \{1,\cdots,J'\}} 1\right] \cdot V\left(\{x_j\}_{j=1,\cdots,J'}\right) \quad (\because\ a(x_j) = 1 \text{ for } x_j \in \Omega,\ b(0_I) = 1) \\
= & V\left(\{x_j\}_{j=1,\cdots,J'}\right).
\end{aligned}$$

□

A.1.1 Proof of Proposition 1

Proof. First, by Tierze extension theorem, we can take a function such that $V : \cup_{J'=1}^{J} \left[\prod_{i=1}^{I} [\min \Omega_i, \max \Omega_i]\right]^{J'} \to \mathbb{R}$. If $\min \Omega > 0$, by Lemma 1, we can construct a continuous function $\overline{V} : \mathcal{M}(I, J; \widetilde{\Omega})\ (\subset \mathbb{R}^{I \times J}) \to \mathbb{R}$ such that $\overline{V}\left(x_1, \cdots, x_{J'}, \underbrace{0_I, \cdots, 0_I}_{J-J'}\right) = V\left(\{x_j\}_{j=1,\cdots,J'}\right)$, for the function $V : \Omega^{\leq J} \to \mathbb{R}$. Let $\overline{V} = V$ if condition (b) holds.

By Corollary 1, there exists a continuous function and $\widetilde{\Lambda} : \left\{\sum_{j=1}^{J} \eta_{I,J}(x_j) | x_j \in \widetilde{\Omega} \subset \mathbb{R}^I\right\} \to Im(\Lambda)\ (\subset \mathbb{R}^J)$ such that $\left\{x_1, \cdots, x_{J'}, \underbrace{0_I, \cdots, 0_I}_{J-J'}\right\} = \widetilde{\Lambda}\left(\sum_{j=1}^{J'} \eta_{I,J}(x_j)\right)$. Hence, by defining $\psi \equiv \overline{V} \circ \widetilde{\Lambda}$, we obtain:

$$\begin{aligned}
V(\{x_1, \cdots, x_{J'}\}) &= \overline{V}\left(x_1, \cdots, x_{J'}, \underbrace{0_I, \cdots, 0_I}_{J-J'}\right) \\
&= \overline{V}\left(\widetilde{\Lambda}\left(\sum_{j=1}^{J'} \eta_{I,J}(x_j)\right)\right) = \psi\left(\sum_{j=1}^{J'} \eta_{I,J}(x_j)\right).
\end{aligned}$$

□

A.1.2 Proof of Proposition 2

Proof. As in the proof of Proposition 1, we can construct a continuous function $\overline{V} : \Omega \times \mathcal{M}(I, J-1; \widetilde{\Omega}) \times \Upsilon \to \mathbb{R}$ such that $\overline{V}\left(x_j, \{x_k\}_{k \in \mathcal{J}-\{j\}}, y\right) = V\left(x_j, \{x_k\}_{k \in \mathcal{J}-\{j\} \ s.t. \ x_k \neq 0}, y\right)$, for the function $V : \Omega \times \Omega^{\leq J-1} \times \Upsilon \to \mathbb{R}$. Hence,

$$\begin{aligned} V\left(x_j, \{x_j\}_{k \in \mathcal{J}-\{j\}}, y\right) &= \overline{V}\left(x_j, \left(\{x_k\}_{k \in \mathcal{J}-\{j\}}, \underbrace{0, \cdots, 0}_{J-|\mathcal{J}|}\right), y\right) \\ &= \overline{V}\left(x_j, \exists\widetilde{\Lambda}\left(\sum_{k \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_k)\right), y\right) \quad (\because \text{Corollary 1}) \\ &= \exists\psi\left(x_j, \sum_{k \in \mathcal{J}-\{j\}} \eta_{I,J-1}(x_k), y\right). \end{aligned}$$

□

A.2 Proof of Proposition 3

Proof. By Proposition 2, we can take functions $\widetilde{g_j} : \mathcal{A}_j \times \mathbb{R}^K \to \mathbb{R}$ and $\widetilde{h_j} : \mathcal{A}_j \times \{x_j^0\} \to \mathbb{R}^K$ such that $\widetilde{\widetilde{g_j}}\left(a_j, \left\{a_{j'}, x_{j'}^0\right\}_{j' \in \mathcal{J}-\{j\}}\right) = \widetilde{g_j}\left(a_j, \sum_{j' \in \mathcal{J}-\{j\}} \widetilde{h_{j'}}(a_{j'}, x_{j'}^0)\right)$. By defining $h_{j'} : \mathcal{A}_j \to \mathbb{R}^K$ such that $h_{j'} \equiv \widetilde{h_{j'}}\left(a_{j'}, x_{j'}^0\right)$, we obtain the statement. □

B Additional results

In this section, we further discuss the implications of the results to other economic models.

B.1 Models with dynamic demand

In the models with dynamic demand, firms need to keep track of many state variables, such as the distribution of heterogeneous consumers' inventory in durable goods, to optimally make pricing and investment decisions (e.g., Goettler and Gordon, 2011). Our results also give insights into such models.

Here, let B be K-dimensional state variables firms need to keep track of. For simplicity, consider the case of a monopolistic firm, and suppose we want to approximate the firm's value functions $V(B)$. When K is large and we don't use any knowledge of the structure of V, in general it is difficult to solve the high-dimensional model. Though it seems no symmetric structure exists in the function V, we can find it based on the knowledge of model structures.

Suppose k-th state variable $B^{(k)}$ is parameterized by n-dimensional parameters $\theta^{(k)}$, and the value function can be rewritten as $V\left(\left\{\left(B^{(k)}, \theta^{(k)}\right)\right\}_{k=1,\cdots,K}\right)$. Here, the function is permutation invariant with respect to $\left(B^{(k)}, \theta^{(k)}\right)$. For instance, consider the model of durable goods, where a monopolistic firm keeps track of discrete type consumers' product holdings. In this case, state variables correspond to the fraction of each type of consumers for each age of the product, and they can be parameterized by the preference parameters θ_{pref} and the age of old products θ_{age}. It is plausible to assume that the order of $\left(B^{(k)}, \theta_{pref}^{(k)}, \theta_{age}^{(k)}\right)$ does not matter.

Then, Proposition 1 implies V can be reformulated as $V\left(\left\{\left(B^{(k)}, \theta^{(k)}\right)\right\}_{k=1,\cdots,K}\right) = \exists\psi\left(\sum_{k=1}^{K} \eta_{1+n,K}\left(B^{(k)}, \theta^{(k)}\right)\right)$, where $\psi : \mathbb{R}^{\kappa(1+n,K)} \to \mathbb{R}$. Hence, we can alternatively use moments $\eta_{1+n,K}\left(B^{(k)}, \theta^{(k)}\right)$ as states.

B.2 Models with Multi-product firms

Though the results in Section 3 cannot be directly applied to the models with multi-product firms, we can further extend the results. The following proposition is relevant to the models with multi-product firms.

Proposition 5. *(Nested structure)*

Let $\Omega \subset [0,\infty)^I$ be a compact subset of $\mathbb{R}^I$, and let Υ be a subset of $\mathbb{R}^C$. Suppose either of the following conditions holds for a continuous permutation invariant function $V : \Omega^{\leq J} \times \left(\Omega^{\leq J}\right)^{\leq (F-1)} \times \Upsilon \to \mathbb{R}$:

(a). $\min \Omega > 0_I$

(b). $V\left(\{x_j\}_{j\in\mathcal{J}_f}, \left\{\{x_j\}_{j\in\mathcal{J}_{\tilde{f}}}\right\}_{\tilde{f}\in\mathcal{F}-\{f\}}, y\right) = V\left(\{x_j\}_{j\in\mathcal{J}_f\ s.t.\ x_j\neq 0}, \left\{\{x_j\}_{j\in\mathcal{J}_{\tilde{f}}\ s.t.\ x_j\neq 0}\right\}_{\tilde{f}\in\mathcal{F}-\{f\}\ s.t.\ \neg(x_j=0\ \forall j\in\mathcal{J}_{f'})}, y\right)$

Then, there exist unique continuous functions $\psi_1 : \mathbb{R}^{\kappa(I,J)} \to \mathbb{R}^{\kappa(IJ,F-1)}$ and $\psi_2 : \mathbb{R}^{\kappa(I,J)} \times \mathbb{R}^{\kappa(IJ,F-1)} \times \mathbb{R}^C \to \mathbb{R}$ such that

$$V\left(\{x_j\}_{j\in\mathcal{J}_f}, \left\{\{x_j\}_{j\in\mathcal{J}_{\tilde{f}}}\right\}_{\tilde{f}\in\mathcal{F}-\{f\}}, y\right) = \psi_2\left(\sum_{k\in\mathcal{J}_f}\eta_{I,J}(x_k), \sum_{\tilde{f}\in\mathcal{F}-\{f\}}\psi_1\left(\sum_{k\in\mathcal{J}_{\tilde{f}}}\eta_{I,J}(x_k)\right), y\right).$$

The proof is shown at the end of this subsection.

Here, consider the dynamic model where each firm decides whether to introduce each product in each period, as in Sweeting (2013). Let x_j be product j's states, namely, whether the product is already introduced at the beginning of the period. We assume that firms follow symmetric MPE, and firm f's value function is permutation invariant with respect to the order of products of the same firms, and to the order of competitors.

The proposition implies that firm f's value function can be reformulated as:

$$V\left(\{x_j\}_{j\in\mathcal{J}_f}, \left\{\{x_j\}_{j\in\mathcal{J}_{\tilde{f}}}\right\}_{\tilde{f}\in\mathcal{F}-\{f\}}\right) = \exists\psi_2\left(\sum_{k\in\mathcal{J}_f}\eta_{I,J}(x_k), \sum_{\tilde{f}\in\mathcal{F}-\{f\}}\exists\psi_1\left(\sum_{k\in\mathcal{J}_{\tilde{f}}}\eta_{I,J}(x_k)\right)\right).$$

. It indicates that value function can be represented as a function of the sum of each rival firm's products' summary statistics $\left(\sum_{\tilde{f}\in\mathcal{F}-\{f\}}\psi_1\left(\sum_{k\in\mathcal{J}_{\tilde{f}}}\eta_{I,J}(x_k)\right)\right)$ and the sum of the moments of own firm's products' states $\left(\sum_{k\in\mathcal{J}_f}\eta_{I,J}(x_k)\right)$.

Proof of Proposition 5

Proof. First, for a $I \times J \times F$ dimensional array A, let $A_f \equiv A[:,:,f] \subset \mathbb{R}^{I\times J}$ and $a_{jf} \equiv A[:,j,f]$. We further define $A_{-f} \subset \mathbb{R}^{I\times J\times(F-1)}$ which corresponds to A skipping A_f. Analogous to the case of two dimensional matrices, let $\mathcal{M}(I,J,F;W) \subset M(I,J,F:\mathbb{R})$ be the set of $I\times J\times F$ dimensional arrays whose column vectors ($A[:,j,f]$ $j=1,\cdots,J,\ f=1,\cdots,F$) are in $W\subset\mathbb{R}^I$, and they are sorted based on the descending lexicographical order.

Then, by Corollary 1, there exists a unique homeomorphism $\Psi_1 : \left\{\sum_{j=1}^J \eta_{I,J}(a_{jf}) | a_{jf}\in\widetilde{\Omega}\subset\mathbb{R}^I\right\} \to Im(\Psi_1)$ $(\subset \mathbb{R}^{I\times J})$ such that:

$$A_f = \Psi_1\left(\sum_{j=1}^J \eta_{I,J}(a_{jf})\right)\ \ \forall f, A_f\in\mathcal{M}(I,J;\widetilde{\Omega}). \tag{1}$$

Furthermore, there exists a unique homeomorphism $\Psi_2 : \left\{\sum_{f'\in\{1,\cdots,F\}-\{f\}}\eta_{IJ,F-1}(A_{f'}) | A_{f'}\in\widetilde{\Omega}^J\subset\mathbb{R}^{I\times J}\right\} \to Im(\Psi_2)$ $(\subset\mathbb{R}^{I\times J\times(F-1)})$ such that:

$$A_{-f} = \Psi_2\left(\sum_{f'\in\{1,\cdots,F\}-\{f\}}\eta_{IJ,F-1}(A_{f'})\right)\ \ \forall A_{-f}\in\mathcal{M}(I,J,F-1;\widetilde{\Omega}). \tag{2}$$

As in the case of Proposition 1, we can construct a continuous permutation invariant function $\overline{V}$: $\mathcal{M}(I,J;\widetilde{\Omega}) \times \mathcal{M}(I,J,F-1;\widetilde{\Omega}) \times \Upsilon \to \mathbb{R}$ such that $\overline{V}(A_f, A_{-f}, y) = V\left(\{x_j\}_{j\in\mathcal{J}_f}, \left\{\{x_j\}_{j\in\mathcal{J}_{\tilde{f}}}\right\}_{\tilde{f}\in\mathcal{F}-\{f\}}, y\right)$ for the function $V : \Omega^{\leq J} \times \left(\Omega^{\leq J}\right)^{\leq(F-1)} \times \Upsilon \to \mathbb{R}$. Here, A_f is a matrix where $\left\{x_{j\in\mathcal{J}_f}, \underbrace{0_I, \cdots, 0_I}_{J-|\mathcal{J}|}\right\}$ are sorted based on the descending lexicographical order, and A_{-f} is a matrix where $\left\{A_{\tilde{f}\in\mathcal{F}-\{f\}}, \underbrace{0_{I\times J}, \cdots, 0_{I\times J}}_{F-|\mathcal{F}|}\right\}$ are sorted based on the descending lexicographical order.

Using equations (1) and (2), we obtain:

$$
\begin{aligned}
&\overline{V}(A_f, A_{-f}, y) \\
= \ &\overline{V}\left(\Psi_1\left(\sum_{j=1}^{J}\eta_{I,J}(a_{jf})\right), \Psi_2\left(\sum_{f'\in\{1,\cdots,F\}-\{f\}}\eta_{IJ,F-1}\left(\Psi_1\left(\sum_{j=1}^{J}\eta_{I,J}(a_{jf'})\right)\right)\right), y\right) \\
= \ &\exists\psi_2\left(\sum_{j=1}^{J}\eta_{I,J}(a_{jf}), \sum_{f'\in\{1,\cdots,F\}-\{f\}}\exists\psi_1\left(\sum_{j=1}^{J}\eta_{I,J}(a_{jf'})\right), y\right).
\end{aligned}
$$

Hence, by the relation between V and $\overline{V}$, we obtain the statement. $\square$