Treatment Effects in Equilibrium Transition

Peilin Yang

Preliminary Proposal

Abstract

Consider the following setting: A policymaker wants to execute a large-scale policy. Before its implementation, we only have a pilot set. Under the SUTVA assumption, we can obtain a ballpark estimation of the policy effect. However, in many large-scale field experiments in economics, we observe strong *general equilibrium* effects, such as significant changes in the price of goods or interest rates, which affect the decisions of the entire population. This proposal tries to address this issue and attempts to find a way to solve the problem. Under a class of data-generating processes commonly used in macroeconomics and industrial organization theory, the final objectives of the research are to (1) provide an unbiased estimator for the global treatment effect and (2) analyze how the bias evolves over time.

1 Motivation

This section discusses the motivation behind providing a new estimator that accounts for the general equilibrium effect.

Typically, general equilibrium effects are not considered in laboratory experiments, or in cases where the treatment does not significantly influence others' decisions. However, this is not the case in a macroeconomic setting. For instance, the Economic Stimulus Act of 2008 in the U.S. involved a \$100 billion program that provided tax rebates to approximately 130 million tax filers. Parker et al. [2013] and Broda and Parker [2014] estimate the treatment effect on the marginal propensity to consume (MPC) as a result of the 2008 tax rebates.[1]

In Johari et al. [2022], the authors describe a two-sided marketing environment involving demand-side 'customer' randomization (CR) and supply-side 'listing' randomization (LR). Similarly, in the setting of Munro et al. [2021], a multiple goods market with a price vector is considered, where consumers make choices among these goods. However, this is not applicable to many economic settings. In a macroeconomic environment, there is typically a single large indivisible market, such as the credit market. In this market, individuals provide savings from the supply side while firms borrow money. The price in this market

[1] This treatment can be viewed as a random treatment for the following reasons. The rebate was disbursed using two methods: either via direct deposit to a bank account, if known by the IRS, or with a mailed paper check. The timing of the disbursement depended on the second-to-last digit of the taxpayer's social security number (SSN). This number provides a source of quasi-experimental variation because the last four digits of a SSN are assigned sequentially to applicants within geographic areas.

influences the distribution of agents, which can be viewed as a covariate shift. We often use the Aiyagari-Bewley-Huggett type models to describe the dynamics of individual behavioral responses to treatments.

One concrete example is Achdou et al. [2022], the treated consumers make decision on consumption c_t and saving a_t to maximize the life-time utility:

$$\mathbf{E_0}\int_0^\infty e^{-\rho t}u(c_t)$$

subject to ODE of the state variable $a_t \in [0,\infty)$

$$\dot{a}_t = y_t + r_t a_t - c_t$$

where r_t is the interest rate and y_t is exogenous random income. The randomness of the covariate arises from the state of efficiency[2]. In particular, we assume that income follows a two-state Poisson process $y_t \in \{y_1, y_2\}$, with $y_2 > y_1$. The process jumps from state 1 to state 2 with intensity λ_1 and from state 2 to state 1 with intensity λ_2. The two states can be interpreted as employment and unemployment, so that λ_1 is the job-finding rate and λ_2 is the job destruction rate. The demand for assets by households is exogenous.

$$D(r_t) = \int_0^\infty a dG_t(a,1) + \int_0^\infty a dG_t(a,2)$$

where $G_t(a,1)$ is the distribution of state-1 agents and $G_t(a,2)$ is the distribution of state-2 agents. Finally, this system can be described by the following *mean-field game* system, similar to the discussion in Achdou et al. [2022] and Light and Weintraub [2022].

By solving the household problem, we can derive the following PDE equations: the first is the Hamilton-Jacobi-Bellman (HJB) equation, and the second is the Kolmogorov equation describing the distribution dynamics:

$$\begin{aligned}
\rho v_j(a,t) &= \max_c u(c) + \partial_a v_j(a,t)\left(y_j + r(t)a - c\right) + \lambda_j\left(v_{-j}(a,t) - v_j(a,t)\right) + \partial_t v_j(a,t),\\
\partial_t g_j(a,t) &= -\partial_a\left[s_j(a,t)g_j(a,t)\right] - \lambda_j g_j(a,t) + \lambda_{-j}g_{-j}(a,t),\\
s_j(a,t) &= y_j + r(t)a - c_j(a,t), \quad c_j(a,t) = (u')^{-1}\left(\partial_a v_j(a,t)\right),
\end{aligned}$$

The difference is that they only discuss the stable state when these variables converge as $t \to \infty$. If we can better understand this system, for example, by determining the transition path between the initial distribution and the final state, we can uncover the underlying error in the estimators we construct. More background about the nature of the system will be discussed in Section 5.

This structure could potentially be extended to experimental settings, especially for large-scale field experiments where the general equilibrium effects cannot be neglected (e.g., Breza and Kinnan [2021], Egger et al. [2022], and Muralidharan et al. [2023]). These papers emphasize the importance of general equilibrium effects in large-scale field experiments. One possible channel is through the variable y_t, which can be viewed as a **treatment** in the experiment, taking either a continuous or dichotomous form. The treatment itself will influence

[2] In the absence of stochastic efficiency, the economy would converge to a mass point.

others' decisions, thus violating the Stable Unit Treatment Value Assumption (SUTVA). A large-scale experiment will affect prices in the economy, which in turn will affect the covariate (savings in the aforementioned model) of the economy. This article aims to discuss this topic and seeks to provide a reasonable evaluation of the policy. The aforementioned model can be considered the underlying basis for analyzing the bias of the estimator.

2 Statistical Model Framework

Now, we turn to statistical modeling. This framework is mainly inspired by Munro et al. [2021] and Johari et al. [2022], but with new elements added.

There are N individuals grouped into control and treatment groups by the vector $\mathbf{W} = \{W_i\}_{i=1}^{N}$. In the first stage, the policymaker will execute a pilot test on a specific, representative population of size $n < N$. The final target is to estimate the total treatment effect when all individuals are treated.

We will make the following assumptions for analyzing the statistical model.

Assumption 1. An individual makes decisions by solving the following dynamic programming problem:

$$V_t(X_t; P_t) = u(Y_t) + \beta \mathbf{E_t}[V_{t+1}(X_{t+1}; P_{t+1})] \tag{1}$$

subject to the constraint

$$c(X_t, X_{t+1}, Y_t; P_t) = 0 \tag{2}$$

where the covariate $X_t \in [\underline{x}, \bar{x}]$. The optimal policy function

$$X_{t+1} = g(X_t, P_t) \tag{3}$$

and

$$Y_t = f(X_t, X_{t+1}, P_t) \tag{4}$$

exist and are unique under some canonical assumptions.[3]

Remark 1. Here, we do not use a continuous optimal control specification for ease of discussion, since the data are discrete, period by period, in the real world. In essence, there is no difference between the discrete and continuous versions, but the latter's uniqueness and existence are provided under some special cases. We will discuss this point later.

Next, by following assumption we get the price mechanism.

Assumption 2. Excessive demand function on period t is

$$Z_{Nt}(P) = D(P_t) - \sum_{i=1}^{N} X_{it}(W_i, X_{it-1}, P_t) \tag{5}$$

There exists a sequence a_n such that $a_N = o(1/\sqrt{N})$.

$$\mathcal{S}_{\mathcal{W}} = \{p : |Z_{Nt}(p)| \leq a_N\}$$

with probability at least $1 - e^{-cN}$ for all N. And $P \in \mathcal{S}_{\mathcal{W}}$.

Remark 2. Under equilibrium, covariate distribution won't change in aggregate even though $X_{it} \neq X_{it-1}$ almost surely.

[3]See the related dynamic programming chapter in Acemoglu [2008].

3 Global Treatment Effect

In this section, we discuss the statistic of interest that we aim to estimate—the global treatment effect (GTE). It is defined as follows:

$$\begin{aligned}\tau_{GTE} = &\sum_{i=1}^{N} E_{X_T|\boldsymbol{W}=\mathbf{1}_N}\left[Y_{Ti}\left(W_i, X_{Ti}, X_{T+1i}, P_T\left(\boldsymbol{W}=\mathbf{1}_N\right)\right)\right] \\ -&\sum_{i=1}^{N} E_{X_1|\boldsymbol{W}=\mathbf{0}_N}\left[Y_{1i}\left(W_i, X_{1i}, X_{2i}, P_1\left(\boldsymbol{W}=\mathbf{0}_N\right)\right)\right]\end{aligned} \tag{6}$$

Our final target is to estimate this after executing a small-scale experiment.

The first estimator we should consider is the mean-in-difference estimator:

$$\hat{\tau}_{GTE}^{D} = \sum_{i=1}^{N} \frac{W_i Y_{Ti}\left(\cdot; P_T\left(\boldsymbol{W}_n\right)\right)}{\eta} - \sum_{i=1}^{N} Y_{1\mathrm{i}}\left(\cdot; P_1\left(\boldsymbol{W}=\mathbf{0}_N\right)\right)$$

where $\eta = n/N$, there are n items in the vector $\mathbf{W}_n$ are 0.

Proposition 3. *$\hat{\tau}_{GTE}^{D}$ is a biased estimator under the setting of Assumption 1 and 2.*

Proof. From the definition of mean-in-difference estimator

$$\begin{aligned}\mathbb{E}\left[\hat{\tau}_{GTE}^{D}\right] = \frac{n}{\eta} \int_{X'_{Ti}} Y_{Ti}\left(1, X'_T, X'_{T+1i}, P_T\left(\boldsymbol{W}_n\right)\right) dP\left(X'_T\right) - \\ N \int_{X_1} Y_{1i}\left(0, X_{1i}, X_{2i}; P_1\left(\boldsymbol{W}=\mathbf{0}_N\right)\right) dP\left(X_1\right)\end{aligned}$$

The deviation

$$\begin{aligned}\tau_{GTE} - \mathbb{E}\left[\hat{\tau}_{GTE}^{D}\right] =& N \int_{X_T} Y_{Ti}\left(1, X_{Ti}, X_{T+1i}; P_T\left(\boldsymbol{W}=\mathbf{1}_N\right)\right) dP\left(X_T\right) \\ &- N \int_{X'_T} Y_{Ti}\left(1, X'_{Ti}, X'_{T+1i}; P_T\left(\boldsymbol{W}_n\right)\right) dP\left(X'_T\right)\end{aligned}$$

$\tau_{GTE} - \mathbb{E}\left[\hat{\tau}_{GTE}^{D}\right] \neq 0$ apparently if covariate X_T shifts and policy function $Y_T(\cdot; p)$ changed with price.

Now let's decomposed the source of bias. If density function f_x exist

$$
\begin{aligned}
&\tau_{GTE} - \mathbb{E}\left[\hat{\tau}^D_{GTE}\right] = N \int_{X_T} Y_{Ti}\left(1, X_{Ti}, X_{T+1i}; P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right) f\left(X_T\right) dX_T \\
&- N \int_{X'_T} Y_{Ti}\left(1, X'_T, X'_{T+1i}; P_T\left(\boldsymbol{W}_n\right)\right) f\left(X'_T\right) dX_T \\
&= N \int_{X_T} \left[Y_{Ti}\left(1, X_{Ti}, g\left(X_{Ti}, P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right), P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right) f(X_T)\right. \\
&\left.- Y_{Ti}\left(1, X'_{Ti}, g\left(X'_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right), P_T\left(\boldsymbol{W}_n\right)\right) f\left(X'_T\right)\right] dX_T \\
&= N \int_{X_T} \left[Y_{Ti}\left(1, X_{Ti}, g\left(X_{Ti}, P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right), P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right) f\left(X_T\right)\right. \\
&- Y_{Ti}\left(1, X_{Ti}, g\left(X_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right), P_T\left(\boldsymbol{W}_n\right)\right) f\left(X_T\right) \\
&+ Y_{Ti}\left(1, X_{Ti}, g\left(X_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right), P_T\left(\boldsymbol{W}_n\right)\right) f\left(X_T\right) \\
&\left.- Y_{Ti}\left(1, X'_{Ti}, g\left(X'_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right), P_T\left(\boldsymbol{W}_n\right)\right) f\left(X'_T\right)\right] dX_T \\
&\approx N \int_{X_T} \left[\left(\frac{\partial Y_{Ti}}{\partial X_{Ti}} \left(g\left(X_{Ti}, P_T\left(\boldsymbol{W} = \mathbf{1}_N\right)\right) - g\left(X_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right)\right)\right.\right. \\
&\left. + \frac{\partial Y_{Ti}}{\partial P_T} \left(P_T\left(\boldsymbol{W} = \mathbf{1}_N\right) - P_T\left(\boldsymbol{W}_n\right)\right)\right) f\left(X_T\right) \\
&\left. + Y_{Ti}\left(1, X_{Ti}, g\left(X_{Ti}, P_T\left(\boldsymbol{W}_n\right)\right), P_T\left(\boldsymbol{W}_n\right)\right)\left(f\left(X_T\right) - f\left(X'_T\right)\right)\right] dX_T
\end{aligned}
$$

These three terms indicate the bias induced by the effect of the policy function change due to price, the direct effect of equilibrium price, and the distribution effect.

□

4 Multiple-Step Estimator

To solve the bias induced by these channels, we can use a multiple-step estimator related to the Olley-Pakes strategy [Olley and Pakes, 1992] and [Pakes and Olley, 1995], but with an additional element: the distribution of covariates.

Algorithm 1 presents a possible strategy that leverages the nature of the underlying data-generating process. Similar to the Olley-Pakes strategy, we can use nonparametric methods to estimate the policy function. If we obtain an accurate estimation of the policy function, we can also derive the corresponding Markov transition matrix of the covariate state X_{it}. By combining these elements, we can ultimately estimate the global treatment effect.

The proof of the asymptotic properties of this multiple-step estimator is crucial. Pakes and Olley [1995] derive an expression for the variance matrix. However, their derivation does not correctly address the problem, as they ignore the variability of the conditional expectation. As a result, their asymptotic variance formula is incorrect (see Hahn et al. [2023]).

Algorithm 1 Multiple-Step Algorithm for GTE

Input: Data Samples from pilot set $\{Y_{ti}, X_{ti}, P_t\}_{i=1}^{N}$.

Procedures

Step 1. Discretize the covariate space $\{x_i\}_{i=1}^{h}$; Estimate the Markov matrix $\widehat{M}(p) = \left[\widehat{m}_{ij}\left(x_i, x_j, p\right)\right]_{ij}$ from data;

Step 2. Estimate the total demand $\widehat{D}(p)$ from series $\left\{\sum_{i=1}^{N} X_{ti}, P_t\right\}$ and outcome function $\hat{Y}_{ti}\left(X_{ti}, X_{t+1i}, P_t\right)$;

Step 3. From matrix $\widehat{M}(p)$ to get the estimation of stable covariate distribution $\widehat{\pi}(x)$:

$$\widehat{M}(p)\hat{\pi}(x) = \hat{\pi}(x)$$

Step 4. Get the estimation of price $\hat{P}$ by solving the market equilibrium

$$\widehat{D}(\hat{P}) = x \cdot \hat{\pi}(x)$$

Step 5. Estimate GTE

$$\hat{\tau}_{GTE}^{M} = N\hat{Y}_{ti}(x, \hat{P}) \cdot \hat{\pi}(x) - \sum_{i=1}^{N} Y_{1i}\left(\cdot; P_1\left(\boldsymbol{W} = \mathbf{0}_N\right)\right)$$

Output: $\hat{\tau}_{GTE}^{M}$

5 Time-Varying Bias Analysis

In the current setting, it is also possible to analyze the bias during the evaluation process from the initial state to convergence. To obtain the theoretical expectation path, we need to solve the following PDE system. This system corresponds to a mean-field game system as described by Lasry and Lions [2007], which is typically expressed in this way.

$$\begin{aligned}
\rho v &= H(x, \nabla v, g) + \nu\Delta v + \partial_t v \quad \text{ in } \mathbb{R}^n \times (0, T) \quad (HJB) \\
\partial_t g &= -\operatorname{div}\left(\nabla_p H(x, \nabla v, g) g\right) + \nu\Delta g \quad \text{ in } \mathbb{R}^n \times (0, T) \quad (KF) \\
g(0) &= g_0, \quad v(x, T) = V(x, g(T)) \quad \text{ in } \mathbb{R}^n.
\end{aligned}$$

It describes the dynamic transition path between the initial value (equilibrium) and the final distribution. The existence and uniqueness of solutions under specific settings, such as systems with non-local coupling and second-order local coupling, can be proved (see Ryzhik's notes, 2023)[4].

Even though in most cases there is no analytical solution, we can still solve the model using numerical methods. By finite sampling from the system, we can analyze the deviation and variance over time.

[4]https://math.stanford.edu/~ryzhik/STANFORD/STANF272-23/lecture-notes-stanf272-23.pdf

References

Daron Acemoglu. *Introduction to modern economic growth*. Princeton university press, 2008.

Yves Achdou, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Income and wealth distribution in macroeconomics: A continuous-time approach. *The review of economic studies*, 89(1):45–86, 2022.

Emily Breza and Cynthia Kinnan. Measuring the equilibrium impacts of credit: Evidence from the indian microfinance crisis. *The Quarterly Journal of Economics*, 136(3):1447–1497, 2021.

Christian Broda and Jonathan A Parker. The economic stimulus payments of 2008 and the aggregate demand for consumption. *Journal of Monetary Economics*, 68:S20–S36, 2014.

Dennis Egger, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker. General equilibrium effects of cash transfers: experimental evidence from kenya. *Econometrica*, 90(6):2603–2643, 2022.

Jinyong Hahn, Zhipeng Liao, and Geert Ridder. Identification and the influence function of olley and pakes'(1996) production function estimator. *Econometric Theory*, 39(5):1044–1066, 2023.

Ramesh Johari, Hannah Li, Inessa Liskovich, and Gabriel Y Weintraub. Experimental design in two-sided platforms: An analysis of bias. *Management Science*, 68(10):7069–7089, 2022.

Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

Bar Light and Gabriel Y Weintraub. Mean field equilibrium: uniqueness, existence, and comparative statics. *Operations Research*, 70(1):585–605, 2022.

Evan Munro, Stefan Wager, and Kuang Xu. Treatment effects in market equilibrium. *arXiv preprint arXiv:2109.11647*, 2021.

Karthik Muralidharan, Paul Niehaus, and Sandip Sukhtankar. General equilibrium effects of (improving) public employment programs: Experimental evidence from india. *Econometrica*, 91(4):1261–1295, 2023.

Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry, 1992.

Ariel Pakes and Steven Olley. A limit theorem for a smooth class of semiparametric estimators. *Journal of Econometrics*, 65(1):295–332, 1995.

Jonathan A Parker, Nicholas S Souleles, David S Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, 103(6):2530–2553, 2013.