The Common Trends Restriction and Dynamic Models of Economic Choice: a Reconciliation

Francesco Ruggieri

February 7, 2023

Background

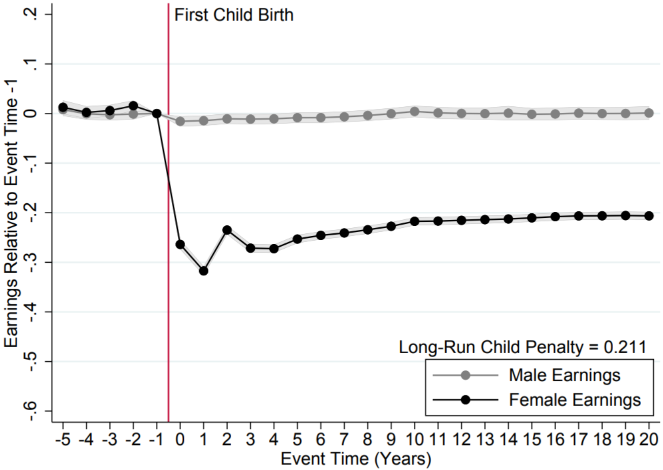
- Difference-in-differences (DiD) designs are widely used for **policy evaluation**
- Recent methodological interest in designs with **staggered adoption** of an absorbing treatment
- This literature has focused on:
  1. The causal interpretation of **linear regression coefficients** under treatment effect heterogeneity
  2. The construction of **alternative estimands** that are immune to the shortcomings of linear regression

Motivation

- Identification in DiD designs hinges on **no anticipation** and **common trends** restrictions
- These assumptions are typically stated within a **dynamic potential outcomes (DPO) model**
- DPO models do **not** require empiricists to specify a **behavioral model** of economic choice
- However, design assumptions in DPOs may **mask** the implied restrictions on **dynamic selection**
- This concern is especially salient if agents **choose** to **sort** into the treated arm...

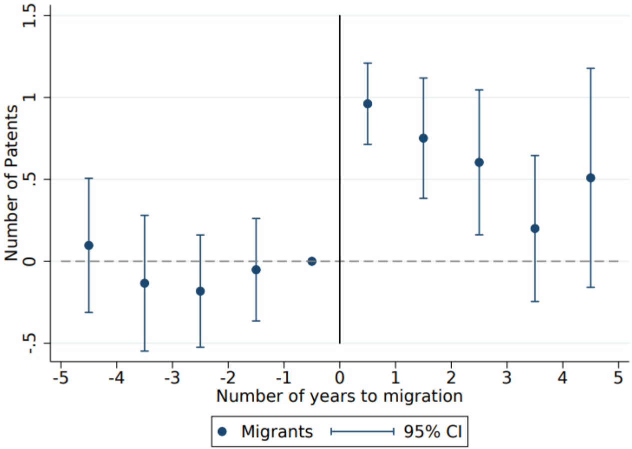
Motivating Example: Kleven, Landais, and Søgaard (2019)

- A staggered DiD design around the time of **child birth** to estimate its effect on **earnings**



Motivating Example: Prato (2022)

- A staggered DiD design around the time of **migration** to estimate its effect on **patenting**



This Discussion

- A recent set of papers investigates the **economic content** of the **common trends** assumption:
  1. *Selection and Parallel Trends* (March 2022), by Ghanem, Sant'Anna, and Wütrich
  2. *Parallel Trends and Dynamic Choices* (July 2022), by Marx, Tamer, and Tang
  3. *Not All Differences-in-Differences Are Equally Compatible with Outcome-Based Selection Models* (October 2022), by de Chaisemartin and d'Haultfœuille
- Each of these papers maps standard DPOs to **economic models** of the outcome
- I will ignore #3 (a short note) and focus on #1, while drawing applications from #2

Setup

- n **units** are indexed by $i \in \{1, \ldots, n\}$ and observed for two **time periods** indexed by $t \in \{1, 2\}$
- $D_{it} \in \{0, 1\}$ indicates unit i's **treatment assignment** at the *beginning* of period t
- $Y_{it} \in \mathbb{R}$ measures unit i's **outcome** at the *end* of period t
- **Sharp design**: the treatment is not available in $t = 1$, i.e., $\mathbb{P}(D_{i1} = 0) = 1$
  - Marx, Tamer, and Tang (2022) considers the richer environment allowed for by fuzzy designs
- One-to-one mapping between **treatment paths** and **time-invariant groups**

$$(D_{i1}, D_{i2}) = (0, 0) \iff G_i = 0 \qquad \text{and} \qquad (D_{i1}, D_{i2}) = (0, 1) \iff G_i = 1$$

Setup

- Following Robins (1986), a **dynamic potential outcomes** model with $Y_{it}(g)$ and $g \in \{0, 1\}$
- A **separable model** for the untreated potential outcome,

$$Y_{it}(0) = A_i + \beta_t + U_{it} \quad \text{with} \quad \mathbb{E}[U_{it}] = 0$$

  The following analysis extends to **nonseparable models** such as $Y_{it}(0) = h_t(A_i, U_{it})$
- A general model of **sorting** into the treated arm,

$$G_i = g(A_i, U_{i1}, U_{i2}, K_i, V_{i1}, V_{i2})$$

  where (K_i, V_{i1}, V_{i2}) are unobserved determinants of the **choice to be treated**

Common Trends and Unrestricted Selection

Assumption (CT): Common Trends in Untreated Potential Outcomes

$Y_{i2}(0) - Y_{i1}(0)$ is mean independent of G_i.

- Let $\mathcal{G}_{\text{all}}$ be the class of **all selection mechanisms** possibly implied by $g(A_i, U_{i1}, U_{i2}, K_i, V_{i1}, V_{i2})$

Proposition 1: Necessary Conditions for (CT) and $g \in \mathcal{G}_{\text{all}}$

Assumption (CT) holds for any $g \in \mathcal{G}_{\text{all}}$ only if $U_{i1} = U_{i2}$ almost surely.

- (CT) is incompatible with both **unrestricted selection** and **time-varying unobservables**
- Because $U_{i1} = U_{i2}$ a.s. is an implausible assumption, it is necessary to **restrict selection**

Common Trends and Restricted Selection

- Consider a **restricted** class of selection mechanisms,

$$\mathcal{G}_1 = \{g \in \mathcal{G}_{\text{all}} : g\left(a, u_1, u_2, k, v_1, v_2\right) = \tilde{g}\left(a, u_1, k, v_1, v_2\right)\}$$

- $\mathcal{G}_1$ restricts sorting **not** to depend on unobserved, time-specific shocks to $Y_{i2}\left(0\right)$

Proposition 2: Necessary Conditions for (CT) and $g \in \mathcal{G}_1$

Assumption (CT) holds for any $g \in \mathcal{G}_1$ only if $\mathbb{E}\left[U_{i2} | A_i, U_{i1}\right] = U_{i1}$ almost surely.

- If selection does not depend on U_{i2}, (CT) is compatible with $\mathbb{P}\left(U_{i1} = U_{i2}\right) \in [0, 1)$

- However, time-varying unobservables must satisfy a **martingale-type** restriction

Common Trends and Further Restricted Selection

- Consider a **further restricted** class of selection mechanisms,

$$\mathcal{G}_2 = \{g \in \mathcal{G}_{\text{all}} : g\,(a, u_1, u_2, k, v_1, v_2) = \tilde{g}\,(a, k, v_1, v_2)\}$$

- $\mathcal{G}_2$ restricts sorting **not** to depend on unobserved, time-specific shocks to $Y_{i1}\,(0)$ and $Y_{i2}\,(0)$

Proposition 3: Necessary Conditions for (CT) and $g \in \mathcal{G}_2$

Assumption (CT) holds for any $g \in \mathcal{G}_2$ only if $\mathbb{E}\,[U_{i2}|A_i] = \mathbb{E}\,[U_{i1}|A_i]$ almost surely.

- If selection does not depend on U_{i1} and U_{i2}, (CT) is compatible with $\mathbb{P}\,(U_{i1} = U_{i2}) \in [0, 1)$

- However, the conditional mean of time-varying unobservables must be **stationary**

Takeaways from Necessary Conditions

- For practically relevant purposes, common trends implies **restrictions** on **sorting** behavior
- **Tighter** restrictions on **selection** allow for **weaker** restrictions on **time-varying unobservables**
- This trade-off illustrates the **economic content** embedded in the common trends assumption

A Two-Period Model of Migration

- To guide the intuition, consider a two-period model of **migration**
  - In $t = 1$, agents live in their home country
  - At the beginning of $t = 2$, they choose whether to **stay** ($G_i = 0$) or **move** ($G_i = 1$)
- Let Y_{it} denote **earnings** and assume that $Y_{it}(0) = A_i + \beta_t + U_{it}$
  - A_i interpretable as the **permanent skill-related** component of earnings
  - β_t interpretable as the **business cycle** component of earnings in the home country

A Two-Period Model of Migration with Selection on the Level

- Consider a choice model that features **selection on the level**
- An agent migrates if lifetime earnings in their home country are **below a subsistence level** c,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[Y_{i1}(0) + \delta Y_{i2}(0) | \mathcal{I}_i\right] \leq c\right]$$

  where $\delta \in [0, 1]$ is a discount factor and $\mathcal{I}_i$ denotes agent i's information set

- Rearranging terms,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[(1 + \delta) A_i + U_{i1} + \delta U_{i2} | \mathcal{I}_i\right] \leq \tilde{c}\right]$$

  with $\tilde{c} \equiv c - \beta_1 - \delta\beta_2$

A Two-Period Model of Migration with Selection on the Gain

- Consider a choice model that features **selection on the gain**, i.e., a Roy model
- Let K_i and V_{i2} denote an individual-specific **migration cost** and **earnings benefit**, respectively
- Migration is a choice described by a simple **dynamic program**:

$$W_{i1} \equiv \mathbb{E}\left[Y_{i1}(0) + \delta \max_{g \in \{0,1\}} \{W_{i2}(g)\} \middle| \mathcal{I}_i\right]$$

with $W_{i2}(0) \equiv Y_{i2}(0)$ and $W_{i2}(1) \equiv Y_{i2}(1) - K_i$

- An individual migrates $(G_i = 1)$ if and only if $\underbrace{\mathbb{E}[V_{i2}|\mathcal{I}_i]}_{\text{expected benefit}} \geq \underbrace{\mathbb{E}[K_i|\mathcal{I}_i]}_{\text{expected cost}}$

Sufficient Conditions for Common Trends in $\mathcal{G}_1$

- A restricted **class of selection mechanisms**,

$$\mathcal{G}_1 = \{g \in \mathcal{G}_{\text{all}} : g\left(a, u_1, u_2, k, v_1, v_2\right) = \tilde{g}\left(a, u_1, k, v_1, v_2\right)\}$$

Proposition 3: Sufficient Conditions for (CT) with $g \in \mathcal{G}_1$

Assumption (CT) holds for any $g \in \mathcal{G}_1$ if

$$\mathbb{E}\left[U_{i2} | A_i, U_{i1}\right] = U_{i1} \;\; \text{a.s.} \quad \text{and} \quad \left(K_i, V_{i1}, V_{i2}\right) | A_i, U_{i1}, U_{i2} \overset{d}{\sim} \left(K_i, V_{i1}, V_{i2}\right) | A_i, U_{i1}$$

- The first condition is also **necessary** for (CT) (Proposition 1)

Sufficient Conditions for Common Trends in $\mathcal{G}_1$

- With selection on the **level**,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[(1+\delta)\,A_i + U_{i1} + \delta U_{i2} | \mathcal{I}_i\right] \leq \tilde{c}\right]$$

If $\mathcal{I}_i = \{A_i, U_{i1}, U_{i2}\}$, (CT) is implied by $\delta = 0$ (full discounting)

- With selection on the **gain**,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[V_{i2}|\mathcal{I}_i\right] \geq \mathbb{E}\left[K_i|\mathcal{I}_i\right]\right]$$

If $\mathcal{I}_i = \{K_i, V_{i2}\}$, (CT) is implied by $(K_i, V_{i2})\,|A_i, U_{i1}, U_{i2} \overset{d}{\sim} (K_i, V_{i2})\,|A_i, U_{i1}$

Sufficient Conditions for Common Trends in $\mathcal{G}_2$

- A further restricted **class of selection mechanisms**,

$$\mathcal{G}_2 = \{g \in \mathcal{G}_{\text{all}} : g\left(a, u_1, u_2, k, v_1, v_2\right) = \tilde{g}\left(a, k, v_1, v_2\right)\}$$

Proposition 4: Sufficient Conditions for (CT) with $g \in \mathcal{G}_2$

Assumption (CT) holds for any $g \in \mathcal{G}_2$ if

$$\mathbb{E}\left[U_{i2}|A_i\right] = \mathbb{E}\left[U_{i1}|A_i\right] \quad \text{a.s.} \quad \text{and} \quad \left(K_i, V_{i1}, V_{i2}\right)|A_i, U_{i1}, U_{i2} \overset{d}{\sim} \left(K_i, V_{i1}, V_{i2}\right)|A_i$$

- The first condition is also **necessary** for (CT) (Proposition 2)

Sufficient Conditions for Common Trends in $\mathcal{G}_2$

- With selection on the **level**,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[(1+\delta)\,A_i + U_{i1} + \delta U_{i2} | \mathcal{I}_i\right] \leq \tilde{c}\right]$$

If $\mathcal{I}_i = \{A_i, U_{i1}, U_{i2}\}$, (CT) is implied by $\delta = 0$ (full discounting) and $U_{i1} = 0$ almost surely

- With selection on the **gain**,

$$G_i \equiv \mathbb{I}\left[\mathbb{E}\left[V_{i2} | \mathcal{I}_i\right] \geq \mathbb{E}\left[K_i | \mathcal{I}_i\right]\right]$$

If $\mathcal{I}_i = \{K_i, V_{i2}\}$, (CT) is implied by $(K_i, V_{i2})\,|A_i, U_{i1}, U_{i2} \overset{d}{\sim} (K_i, V_{i2})\,|A_i$

Takeaways from Sufficient Primitive Conditions

- The **plausibility** of the common trends assumption is **context-specific**
- Before implementing a DiD design, it may be useful to **sketch** a model of economic choice
    - Agents' information set may be particularly salient
- The model can offer guidance on **restrictions** implied by alternative **selection mechanisms**
- This analysis may help determine if (CT) is or is not **compatible** with agents' sorting behavior

Conclusion

- Perhaps unsurprisingly, **DiD designs** and standard **panel data models** are linked
- In practice, the **common trends** assumption **restricts** sorting and/or time-varying unobservables
- Its context-specific plausibility should be assessed based on **economic** (vs. statistical) **arguments**