Online Appendix:
Stage-Based Identification of Policy Effects

Christian Alemán
NYU Abu Dhabi

Christopher Busch
LMU Munich
CESifo

Alexander Ludwig
Goethe University Frankfurt
ICIR, UAB, BSE and CEPR

Raül Santaeulàlia-Llopis
NYU Abu Dhabi
UAB, BSE and CEPR

October 18, 2023

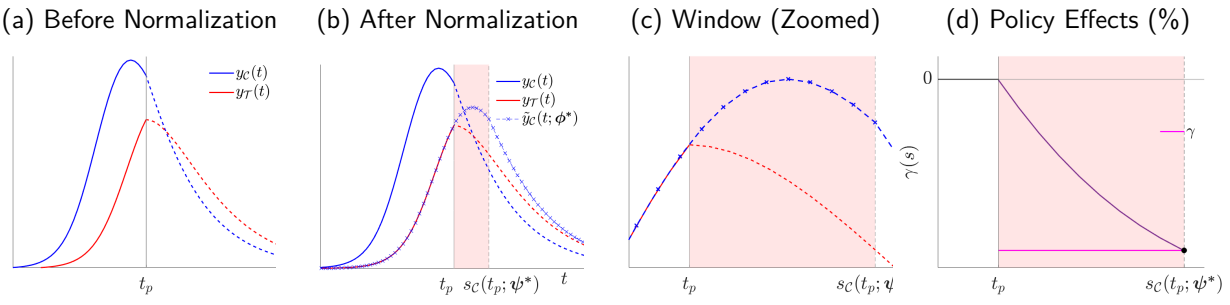
O.A Further Illustrations: Policy After the Peak

Here, within the context of our benchmark illustration in the paper Section 2, we assess additional examples in which the policy is implemented when regional outcome paths have surpassed their peak in one region in Section O.A.1 and in two regions in Section O.A.2.

O.A.1 Policy After the Peak: One Region

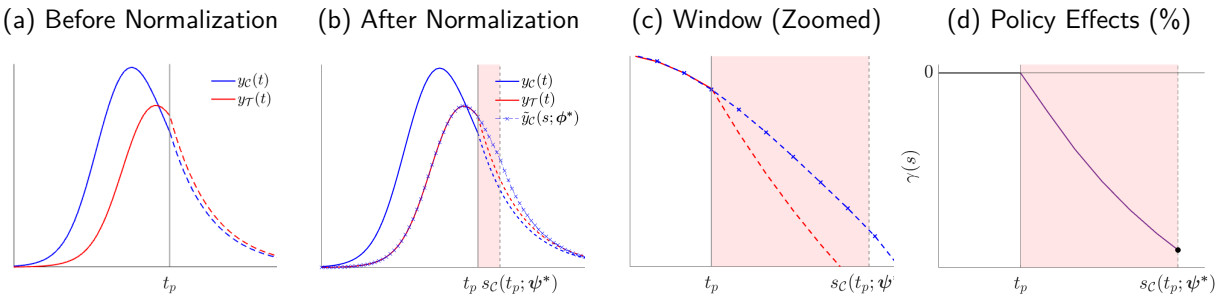
Consider a scenario in which the nationwide policy is implemented at the same time in two regions $r = \{\mathcal{C}, \mathcal{T}\}$. Assume the policy arrives before the outcome path of region $\mathcal{T}$ reaches its peak and, at the same time, after the outcome path of region $\mathcal{C}$ as surpassed its peak; see panel (a) of Figure O.A.1. We pick region $\mathcal{T}$ as reference and apply SBI. That is, we map region $\mathcal{C}$ onto region $\mathcal{T}$ using pre-policy data only. This mapping generates the normalized outcome path $\widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi}^*)$; see panel (b) in Figure O.A.1. Our normalization opens a window in stages between t_p and $s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*)$ in which we identify the effects of policy. Since region $\mathcal{C}$ is at a more advanced stage at the time of policy implementation, i.e., $t_p < s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*)$, it serves as no-policy counterfactual for region $\mathcal{T}$ inside the identification window. We zoom the identification window in panel (c) of Figure O.A.1 and the associated policy effects in panel (d) of Figure O.A.1.

Figure O.A.1: Stage-Based Identification of Policy Effects: Further Illustrations: One Region After the Peak



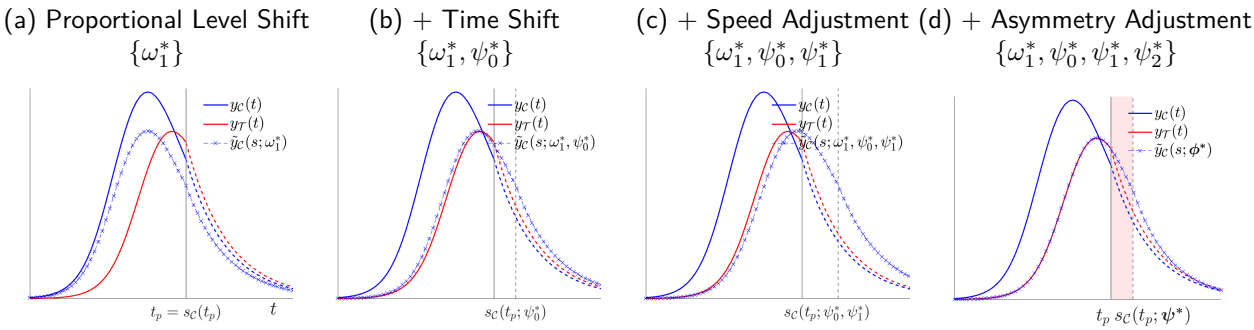
Notes: See the notes in Figure 3.

Figure O.A.2: Stage-Based Identification of Policy Effects: Further Illustrations: Two Regions After the Peak



Notes: See the notes in Figure 3.

Figure O.A.3: Decomposing by Normalization Coefficient: Further Illustrations: Two Regions After the Peak



Notes: See the notes in Figure 4.

O.A.2 Policy After the Peak: Two Regions

Here we discuss a scenario analogous to that of the previous section when the nationwide policy arrives to both regions after their outcome paths has surpassed their respective peaks; see panel (a) of Figure O.A.2. Picking again region $\mathcal{T}$ as reference, we map region $\mathcal{C}$ onto region $\mathcal{T}$ using pre-policy data only, which generates the normalized outcome path $\widetilde{y}_{\mathcal{C}}(t;\phi^*)$; see panel (b) in Figure O.A.2. Hence, SBI opens a window in stages between t_p and $s_{\mathcal{C}}(t;\psi^*)$ in which the effects of policy are identified. We zoom the identification window in panel (c) of Figure O.A.2 and the associated policy effects in panel (d) of Figure O.A.2. In Figure O.A.3, we further unpack the contribution of each normalization coefficient in generating the normalized path. Note that in this scenario, since the outcome paths of the two regions is already affected by the potential asymmetry in which the outcome paths increase before their respective peaks and decrease after the peaks, the additional parameter ψ_2 that asymmetrically shapes stages into time—in $t_{\mathcal{C}} = \psi_0 + \psi s + \psi_2 s^2$—plays a role.

O.B Some Analytical Derivations

Here, we follow our discussion in the paper Section 2.3 and provide analytical derivations for some cases in which we can explicitly express the normalization coefficients as functions of the structural parameters of a known data generating process. As emphasized in our main text, note, again, that this exercise merely serves to illustrate our method and provides an interpretation of the normalization coefficients as those that aim to reshape the structural parameters of the non-reference regions into those of the reference region. Indeed, if the data generating process were actually known, there would be no need to apply SBI; or any other identification method for that matter.

Our method works under the proposition that if there exists a composite function (1) such that (9) holds with equality, then the normalization procedure—the minimization of (5) subject to (2) and (6)—recovers the coefficients $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\}$ up to a minimization error by approximating the functions $t_{\mathcal{C}}(.) \approx t_{\mathcal{C}}(.;\boldsymbol{\psi})$, $f_{\mathcal{C}}(.) \approx f_{\mathcal{C}}(.;\boldsymbol{\omega})$ and, hence, $\widetilde{y}_{\mathcal{C}}(.) \approx \widetilde{y}_{\mathcal{C}}(.;\boldsymbol{\phi})$ for all $s \in \mathbb{C}(s)$. Thus, under the identification assumption, we can identify the policy effects for all $s \in \mathbb{W}(s;\boldsymbol{\psi}^*)$. Here, we are interested in cases where (9) and (2) hold with equality and, hence, analytical solutions for the normalization coefficients $\boldsymbol{\phi}$ potentially exist for all $s \in \mathbb{C}(s)$. For this discussion, we focus on cases in which $t_{\mathcal{C}}$ and $f_{\mathcal{C}}$ are linear—$t_{\mathcal{C}} = \psi_0 + \psi_1 s$ and $f_{\mathcal{C}} = \omega_0 + \omega_1 y_{\mathcal{C}}(t)$, and study time paths, $y_r(t)$, that are described by trigonometric functions (Section O.B.1), polynomial functions (Section O.B.2) and generalized logistic functions (Section O.B.3) under the assumption that we know these data generating processes.

As in the paper Section 2.3, the outcome time paths throughout this Appendix O.B follow:

$$y_r(t) = \left(1 - \gamma_{r,t}\mathbf{1}_{t \geq t_p}\right) g(t;\Theta_r)$$

for regions $r = \{\mathcal{C}, \mathcal{T}\}$ and periods $t \in \{0, \ldots, t_p, \ldots, T\}$. The region-specific policy effects for periods after t_p are captured by $\gamma_{r,t}$ whereas $g(t;\Theta_r)$ captures the outcome time path of region r (determined by region-specific structural parameters Θ_r) occurring absent policy—that is, $g(t;\Theta_r)$ is the true region-specific no-policy counterfactual for $t \geq t_p$.

O.B.1 Trigonometric Functions

Here, we study a case with trigonometric time paths (see panel (b) of Figure 5). Precisely, we define

$$g(t;\Theta_r) = \theta_{1,r} \sin\left(\theta_{3,r} t + \theta_{2,r}\right) + \theta_{0,r},$$

for regions $r = \{\mathcal{C}, \mathcal{T}\}$ and periods $t \in \{0, \ldots, t_p, \ldots, T\}$. Then, picking a region (here: $\mathcal{T}$) as reference, we apply our normalization procedure. That is, we map the outcome path of the non-reference region $\mathcal{C}$ onto that of the reference region $\mathcal{T}$ using pre-policy data only. To do so, we postulate a normalized path for the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = \left(f_{\mathcal{C}}(.;\boldsymbol{\omega}) \circ y_{\mathcal{C}} \circ t_{\mathcal{C}}(.;\boldsymbol{\psi})\right)(s)$ with outer composite $f_{\mathcal{C}} = \omega_0 + \omega_1 y_{\mathcal{C}}(t)$ and inner composite $t_{\mathcal{C}} = \psi_0 + \psi_1 s$. That is,

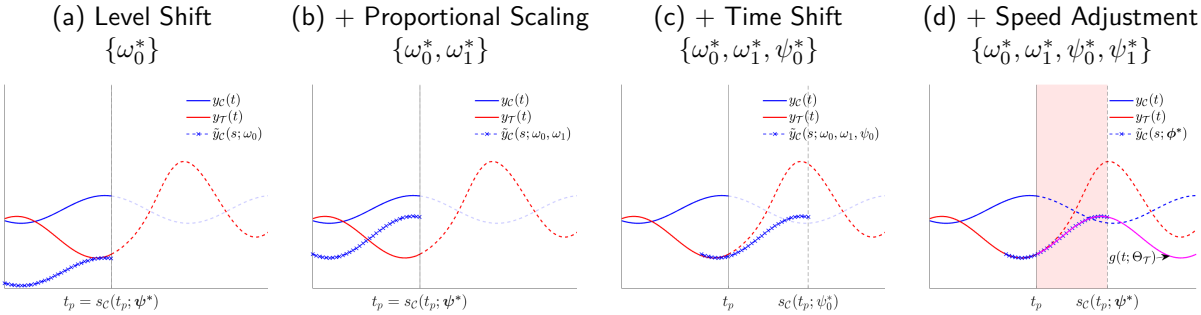
$$\begin{aligned}
\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) &= \omega_1 y_{\mathcal{C}}(\psi_0 + \psi_1 s) + \omega_0 \\
&= \omega_1 \left(\theta_{1,\mathcal{C}} \sin\left(\theta_{3,\mathcal{C}}(\psi_0 + \psi_1 s) + \theta_{2,\mathcal{C}}\right) + \theta_{0,\mathcal{C}}\right) + \omega_0 \\
&= \underbrace{\omega_1 \theta_{1,\mathcal{C}}}_{=\theta_{1,\mathcal{T}}} \sin\left(\underbrace{\theta_{3,\mathcal{C}}\psi_1}_{=\theta_{3,\mathcal{T}}} s + \underbrace{\left(\theta_{3,\mathcal{C}}\psi_0 + \theta_{2,\mathcal{C}}\right)}_{=\theta_{2,\mathcal{T}}}\right) + \underbrace{\omega_1 \theta_{0,\mathcal{C}} + \omega_0}_{=\theta_{0,\mathcal{T}}} = y_{\mathcal{T}}(s),
\end{aligned}$$

where the last equality emerges from holding equation (9) with equality, i.e., $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\psi}) = y_{\mathcal{T}}(s)$ (note that $s = t$ for the reference region $\mathcal{T}$). Then, we find the undetermined normalization coefficients as,

$$\begin{aligned}
\omega_0 &= \theta_{0,\mathcal{T}} - \frac{\theta_{1,\mathcal{T}}}{\theta_{1,\mathcal{C}}}\theta_{0,\mathcal{C}}, \qquad & \omega_1 &= \frac{\theta_{1,\mathcal{T}}}{\theta_{1,\mathcal{C}}}, \\
\psi_0 &= \frac{\theta_{2,\mathcal{T}} - \theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}, & \psi_1 &= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}}.
\end{aligned}$$

That is, the normalization coefficients have an exact and unique solution. For the illustration in panel (b) of Figure 5, we set $\Theta_{\mathcal{C}} = \{1.0, 5.0, 4.1, 1.00\}$ and $\Theta_{\mathcal{T}} = \{1.5, 3.0, 1.1, 1.05\}$, which implies the following unique solution for $\boldsymbol{\phi}^* = \{\psi_0^*, \psi_1^*, \omega_0^*, \omega_1^*\} = \{-3.0, 1.05, -4.5, 1.5\}$. We show the role of each of these normalizaiton coefficients in the non-orthogonal decomposition in Figure O.B.1.

Figure O.B.1: Decomposition by Normalization Coefficient: Trigonometric Time Paths



Notes: See the notes in Figure 4.

O.B.2 Polynomial Functions

Here we use a set of polynomial functions to generate outcome time paths. We discuss a case with unique solutions for $\boldsymbol{\phi}^*$ that emerges from cubic time paths and a case with multiple solutions for $\boldsymbol{\phi}^*$ that emerges from quadratic time paths.

O.B.2.1 Cubic Time Paths

Here, we assume cubic time paths (see panel (c) of Figure 5):

$$g(t;\Theta_r) = \theta_{0,r} + \theta_{1,r}t + \theta_{2,r}t^2 + \theta_{3,r}t^3$$

for regions $r = \{\mathcal{C}, \mathcal{T}\}$ and periods $t \in \{0, \ldots, t_p, \ldots, T\}$. Then, picking a region (here: $\mathcal{T}$) as reference, we apply our normalization procedure. That is, we map the outcome path of the non-reference region $\mathcal{C}$ onto that of the reference region $\mathcal{T}$ using pre-policy data only. To do so, we postulate a normalized path for the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = (f_{\mathcal{C}}(.;\boldsymbol{\omega}) \circ y_{\mathcal{C}} \circ t_{\mathcal{C}}(.;\boldsymbol{\psi}))(s)$ with outer composite $f_{\mathcal{C}} = \omega_0 + \omega_1 y_{\mathcal{C}}(t)$ and inner composite $t_{\mathcal{C}} = \psi_0 + \psi_1 s$. That is,

$$\begin{aligned}
\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) &= \omega_1 y_{\mathcal{C}}(\psi_0 + \psi_1 s) + \omega_0 \\
&= \omega_1 \left(\theta_{0,\mathcal{C}} + \theta_{1,\mathcal{C}}(\psi_0 + \psi_1 s) + \theta_{2,\mathcal{C}}(\psi_0 + \psi_1 s)^2 + \theta_{3,\mathcal{C}}(\psi_0 + \psi_1 s)^3\right) + \omega_0 \\
&= \omega_0 + \omega_1 \left(\theta_{0,\mathcal{C}} + \theta_{1,\mathcal{C}}\psi_0 + \theta_{2,\mathcal{C}}\psi_0^2 + \theta_{3,\mathcal{C}}\psi_0^3\right) \\
&\quad + \omega_1 \left(\theta_{1,\mathcal{C}}\psi_1 + 2\theta_{2,\mathcal{C}}\psi_0\psi_1 + 3\theta_{3,\mathcal{C}}\psi_0^2\psi_1\right) s \\
&\quad + \omega_1 \left(\theta_{2,\mathcal{C}}\psi_1^2 + 3\theta_{3,\mathcal{C}}\psi_0\psi_1^2\right) s^2 \\
&\quad + \omega_1 \theta_{3,\mathcal{C}}\psi_1^3 s^3 = y_{\mathcal{T}}(s)
\end{aligned}$$

where the last equality emerges from holding equation (9) with equality, i.e., $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\psi}) = y_{\mathcal{T}}(s)$ (note that $s = t$ for the reference region $\mathcal{T}$). This implies the following system with four equations and four unknowns, $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\} = \{\psi_0, \psi_1, \omega_0, \omega_1\}$:

$$\theta_{0,\mathcal{T}} = \theta_{0,\mathcal{T}}(\omega_0, \omega_1, \psi_0) = \omega_0 + \omega_1 \left(\theta_{0,\mathcal{C}} + \theta_{1,\mathcal{C}}\psi_0 + \theta_{2,\mathcal{C}}\psi_0^2 + \theta_{3,\mathcal{C}}\psi_0^3\right) \tag{O.B.1}$$

$$\theta_{1,\mathcal{T}} = \theta_{1,\mathcal{T}}(\omega_1, \psi_0, \psi_1) = \omega_1 \left(\theta_{1,\mathcal{C}}\psi_1 + 2\theta_{2,\mathcal{C}}\psi_0\psi_1 + 3\theta_{3,\mathcal{C}}\psi_0^2\psi_1\right) \tag{O.B.2}$$

$$\theta_{2,\mathcal{T}} = \theta_{2,\mathcal{T}}(\omega_1, \psi_0, \psi_1) = \omega_1 \left(\theta_{2,\mathcal{C}}\psi_1^2 + 3\theta_{3,\mathcal{C}}\psi_0\psi_1^2\right) \tag{O.B.3}$$

$$\theta_{3,\mathcal{T}} = \theta_{3,\mathcal{T}}(\omega_1, \psi_1) = \omega_1 \theta_{3,\mathcal{C}}\psi_1^3 \tag{O.B.4}$$

Then, we find the undetermined normalization coefficients ($\boldsymbol{\phi}$) that solve the system (O.B.1)-(O.B.4) using the following steps:

STEP 1. Isolate ω_1 in $\theta_{3,\mathcal{T}}(\omega_1, \psi_1)$,

$$\omega_1 = \omega_1(\psi_1) = \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^3}$$

STEP 2. Plug $\omega_1(\psi_1)$ in $\theta_{3,\mathcal{T}}(\omega_1, \psi_0, \psi_1)$ and isolate ψ_0:

$$\begin{aligned}
\theta_{2,\mathcal{T}} &= \omega_1 \left(\theta_{2,\mathcal{C}}\psi_1^2 + 3\theta_{3,\mathcal{C}}\psi_0\psi_1^2\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^3} \left(\theta_{2,\mathcal{C}}\psi_1^2 + 3\theta_{3,\mathcal{C}}\psi_0\psi_1^2\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1} \left(\theta_{2,\mathcal{C}} + 3\theta_{3,\mathcal{C}}\psi_0\right)
\end{aligned}$$

hence,

$$\psi_0(\psi_1) = \frac{1}{3}\frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\psi_1 - \frac{1}{3}\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}$$

STEP 3. Plug $\omega_1(\psi_1)$ and $\psi_0(\psi_1)$ in $\theta_{1,\mathcal{T}}(\omega_1, \psi_0, \psi_1)$ and isolate ψ_1:

$$\begin{aligned}
\theta_{1,\mathcal{T}} &= \omega_1 \left(\theta_{1,\mathcal{C}}\psi_1 + 2\theta_{2,\mathcal{C}}\psi_0\psi_1 + 3\theta_{3,\mathcal{C}}\psi_0^2\psi_1\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^3} \left(\theta_{1,\mathcal{C}}\psi_1 + 2\theta_{2,\mathcal{C}}\psi_0\psi_1 + 3\theta_{3,\mathcal{C}}\psi_0^2\psi_1\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^2} \left(\theta_{1,\mathcal{C}} + 2\theta_{2,\mathcal{C}}\psi_0 + 3\theta_{3,\mathcal{C}}\psi_0^2\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^2} \left(\theta_{2,\mathcal{C}} + 2\theta_{3,\mathcal{C}} \left(\underbrace{\frac{1}{3}\frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\psi_1 - \frac{1}{3}\frac{\theta_{3,\mathcal{C}}}{\theta_{3,\mathcal{C}}}}_{\psi_0(\psi_1)}\right) + 3\theta_{3,\mathcal{C}} \left(\underbrace{\frac{1}{3}\frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\psi_1 - \frac{1}{3}\frac{\theta_{3,\mathcal{C}}}{\theta_{3,\mathcal{C}}}}_{\psi_0(\psi_1)}\right)^2\right) \\
&= \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}\psi_1^2} \left(\theta_{1,\mathcal{C}} + 2\theta_{2,\mathcal{C}} \left(\frac{1}{3}\frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\psi_1 - \frac{1}{3}\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}\right) + 3\theta_{3,\mathcal{C}} \left(\frac{1}{9}\left(\frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\right)^2\psi_1^2 + \frac{1}{9}\left(\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}\right)^2 - 2\frac{1}{9}\frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}\psi_1\right)\right) \\
&= \frac{1}{3}\frac{\theta_{2,\mathcal{T}}^2}{\theta_{3,\mathcal{T}}} + \frac{2}{3}\left(\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{T}}}\theta_{2,\mathcal{T}} - \frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\theta_{2,\mathcal{C}}\right)\frac{1}{\psi_1} + \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}}\left(\theta_{1,\mathcal{C}} - \frac{1}{3}\frac{\theta_{2,\mathcal{C}}^2}{\theta_{3,\mathcal{C}}}\right)\frac{1}{\psi_1^2}
\end{aligned}$$

That is,

$$\underbrace{\frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}}\left(\theta_{1,\mathcal{C}} - \frac{1}{3}\frac{\theta_{2,\mathcal{C}}^2}{\theta_{3,\mathcal{C}}}\right)}_{a}\frac{1}{\psi_1^2} + \underbrace{\frac{2}{3}\left(\frac{\theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{T}}}\theta_{2,\mathcal{T}} - \frac{\theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{T}}}\theta_{2,\mathcal{C}}\right)}_{b=0}\frac{1}{\psi_1} + \underbrace{\frac{1}{3}\frac{\theta_{2,\mathcal{T}}^2}{\theta_{3,\mathcal{T}}} - \theta_{1,\mathcal{T}}}_{c} = 0$$

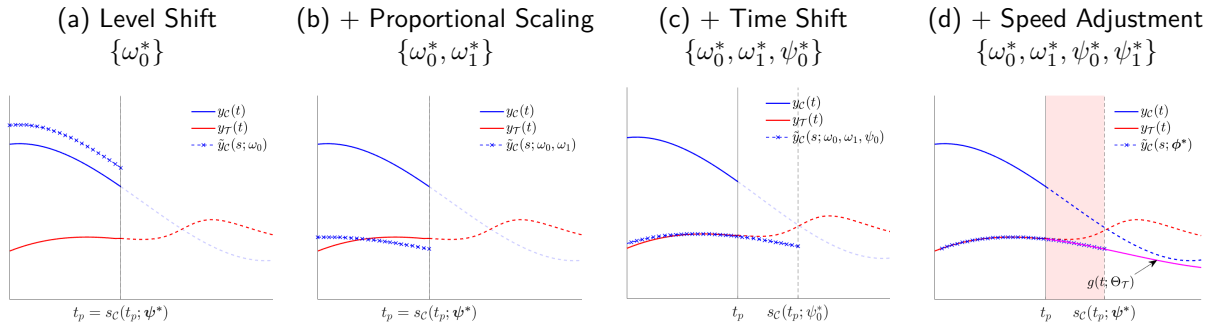
where note that $b = 0$. Thus, we can isolate $\psi_1 = \sqrt{\frac{a}{-c}}$ which delivers a unique solution for ψ_1 if $\frac{a}{-c} > 0$.[1]

STEP 4. Plug ψ_1 into $\omega_1(\psi_1)$ and $\psi_0(\psi_1)$ in order to recover ω_1 and ψ_0.

STEP 5. Plug ω_1, ψ_0 and ψ_1 into $\theta_{0,\mathcal{T}}(\omega_0, \omega_1, \psi_0)$ and isolate (recover) ω_0.

[1]One can further elaborate this to show that $\frac{a}{-c} = \frac{(3\theta_{1,\mathcal{C}}\theta_{3,\mathcal{C}} - \theta_{2,\mathcal{C}}^2)/\theta_{3,\mathcal{C}}^2}{(3\theta_{1,\mathcal{T}}\theta_{3,\mathcal{T}} - \theta_{2,\mathcal{T}}^2)/\theta_{3,\mathcal{T}}^2}$ and hence the solution to the system (O.B.1)-(O.B.4) emerging from the cubic time paths exists as long as the term in the numerator $(3\theta_{1,\mathcal{C}}\theta_{3,\mathcal{C}} - \theta_{2,\mathcal{C}}^2)$ and the term in the denominator $(3\theta_{1,\mathcal{T}}\theta_{3,\mathcal{T}} - \theta_{2,\mathcal{T}}^2)$ have the same sign.

Figure O.B.2: Decomposition by Normalization Coefficient: Cubic Time Paths



Notes: See the notes in Figure 4.

That is, the normalization coefficients have an exact and unique solution. For the illustration in panel (c) of Figure 5, we set $\Theta_{\mathcal{C}} = \{12, -4.5, 0.3, 0.2\}$ and $\Theta_{\mathcal{T}} = \{7.84, -1.28, -0.17, 0.07\}$, which implies the following unique solution for $\boldsymbol{\phi}^* = \{\psi_0^*, \psi_1^*, \omega_0^*, \omega_1^*\} = \{-1.350, 1.093, 2.788, 0.278\}$. We show the role of each of these normalization coefficients in the non-orthogonal decomposition in Figure O.B.2.

O.B.2.2 Quadratic Time Paths

Here, we assume quadratic time paths:

$$g(t;\Theta_r) = \theta_{0,r} + \theta_{1,r}t + \theta_{2,r}t^2$$

for regions $r = \{\mathcal{C}, \mathcal{T}\}$ and periods $t \in \{0, \ldots, t_p, \ldots, T\}$. Then, picking a region (here: $\mathcal{T}$) as reference, we apply the normalization procedure. That is, we map the outcome path of the non-reference region $\mathcal{C}$ onto that of the reference region $\mathcal{T}$ using pre-policy data only. To do so, we postulate a normalized path for the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = (f_{\mathcal{C}}(.;\boldsymbol{\omega}) \circ y_{\mathcal{C}} \circ t_{\mathcal{C}}(.;\boldsymbol{\psi}))(s)$ with outer composite $f_{\mathcal{C}} = \omega_0 + \omega_1 y_{\mathcal{C}}(t)$ and inner composite $t_{\mathcal{C}} = \psi_0 + \psi_1 s$. That is,

$$\begin{aligned}
\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\psi}) &= \omega_1 y_{\mathcal{C}}(\psi_0 + \psi_1 s) + \omega_0 \\
&= \omega_1 \left(\theta_{0,\mathcal{C}} + \theta_{1,\mathcal{C}}(\psi_0 + \psi_1 s) + \theta_{2,\mathcal{C}}(\psi_0 + \psi_1 s)^2\right) + \omega_0 \\
&= \omega_1\theta_{0,\mathcal{C}} + \omega_1\left(\theta_{1,\mathcal{C}}\psi_0 + \theta_{1,\mathcal{C}}\psi_1 s\right) + \omega_1\left(\theta_{2,\mathcal{C}}\psi_0^2 + \theta_{2,\mathcal{C}}2\psi_0\psi_1 s + \theta_{2,\mathcal{C}}\psi_1^2 s^2\right) + \omega_0 \\
&= \omega_1\theta_{0,\mathcal{C}} + \omega_1\theta_{1,\mathcal{C}}\psi_0 + \omega_1\theta_{2,\mathcal{C}}\psi_0^2 + \omega_0 + \left(\omega_1\theta_{1,\mathcal{C}}\psi_1 + \omega_1\theta_{2,\mathcal{C}}2\psi_0\psi_1\right)s + \omega_1\theta_{2,\mathcal{C}}\psi_1^2 s^2 = y_{\mathcal{T}}(s)
\end{aligned}$$

where the last equality emerges from holding equation (9) with equality, i.e., $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\psi}) = y_{\mathcal{T}}(s)$ (note that $s = t$ for the reference region $\mathcal{T}$). This implies the following system with three equations and four unknowns, $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\} = \{\psi_0, \psi_1, \omega_0, \omega_1\}$:

$$\theta_{0,\mathcal{T}} = \theta_{0,\mathcal{T}}(\omega_0, \omega_1, \psi_0) = \omega_0 + \omega_1\left(\theta_{0,\mathcal{C}} + \theta_{1,\mathcal{C}}\psi_0 + \theta_{2,\mathcal{C}}\psi_0^2\right) \tag{O.B.5}$$

$$\theta_{1,\mathcal{T}} = \theta_{1,\mathcal{T}}(\omega_1, \psi_0, \psi_1) = \omega_1\left(\theta_{1,\mathcal{C}}\psi_1 + \theta_{2,\mathcal{C}}2\psi_0\psi_1\right) \tag{O.B.6}$$

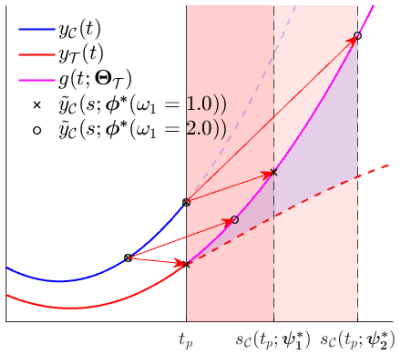
$$\theta_{2,\mathcal{T}} = \theta_{2,\mathcal{T}}(\omega_1, \psi_1) = \omega_1\theta_{2,\mathcal{C}}\psi_1^2 \tag{O.B.7}$$

This is an underidentified system with multiple exact solutions. For example, for any value of ω_1 (or ψ_2) we obtain a different exact solution of $\boldsymbol{\phi}^*$.

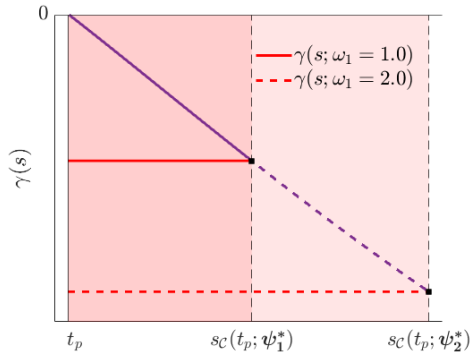
As an illustration, let $\Theta_{\mathcal{C}} = \{8, -5, 3\}$ and $\Theta_{\mathcal{T}} = \{4, -4, -2\}$. Then, initiating the system at $\omega_1 = 1.0$, we obtain the solution $\boldsymbol{\phi}^* = \{\psi_0^*, \psi_1^*, \omega_0^*, \omega_1^*\} = \{0.016, 0.816, -3.916, 1.0\}$. Alternatively, initiating the system at $\omega_1 = 2.0$, we obtain the alternative solution $\boldsymbol{\phi}^* = \{\psi_0^*, \psi_1^*, \omega_0^*, \omega_1^*\} = \{0.256, 0.577, -9.833, 2.0\}$; see panel (a) of Figure O.B.3. Then, as Remark 1 in Section 2.3 states, note that these two (and all for that matter) solutions ($\boldsymbol{\phi}^*$) to the system (O.B.5)-(O.B.7) deliver a counterfactual path that is exactly identical to the true

Figure O.B.3: Quadratic Time Paths

(a) After Normalization



(b) Identified Policy Effects(%)



Notes: This is an example with multiple solutions for $\boldsymbol{\phi}^*$. In panel (a), as an illustration, we show two mappings of $\mathcal{C}$ onto $\mathcal{T}$ (reference region) that depend on two alternative solutions $\boldsymbol{\phi}^*(\omega_1)$ which are a function of a choice for ω_1. In panel (b), we show the (interim) policy effects $\gamma(s)$ which are identical across solutions. We also show the (cumulative) policy effects γ which depend not only on the interim effects, but also on the size of the window and, hence, on the solution $\boldsymbol{\phi}^*(\omega_1)$.

counterfactual path without policy, $y_{\mathcal{T}}(s) = g(s;\Theta_{\mathcal{T}})$. In panel (a) of Figure O.B.3, we graphically show the identified counterfactual path, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})$, of two (of the many possible) solutions assuming either $\omega_1 = 1.0$ (crossed markers) or $\omega_1 = 2.0$ (circle markers). This implies that the (interim) policy effects $\gamma(s)$ are identical across both (all for that matter) solutions of the normalization coefficients ($\boldsymbol{\phi}^*$); see panel (b) of Figure O.B.3. At the same time, note that the overall policy effect γ is determined by the behavior of the interim policy effects and the size of the window, $\mathbb{W}(s,\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*(\omega_1)]$, which differs by the solution. Further, note that if policy is applied non-nationwide (e.g. assume that the policy is never implemented in region $\mathcal{C}$), then the overall policy effect γ is also identical across potential multiple solutions of $\boldsymbol{\phi}^*$ because the identification window is open, i.e., $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, \infty)$.

O.B.3 Generalized Logistic Functions

Here, we assume the outcome time paths follow the log of a generalized logistic function,

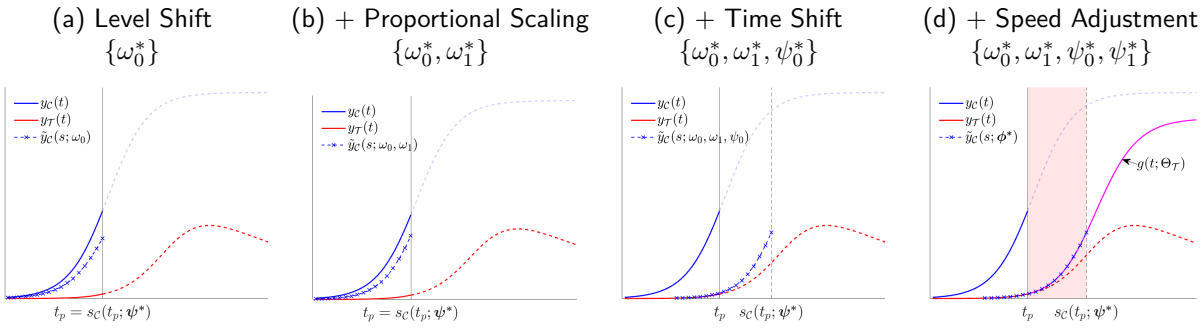
$$g(t;\Theta_r) = \ln\left(\frac{\theta_{1,r}}{\left(1+\exp\left(-\theta_{3,r}t+\theta_{2,r}\right)\right)^{\frac{1}{\theta_{4,r}}}}\right) \tag{O.B.8}$$

for regions $r = \{\mathcal{C},\mathcal{T}\}$ and periods $t \in \{0,\ldots,t_p,\ldots,T\}$. Note that the outcome time path in equation (O.B.8) generalizes the logistic function discussed in the paper Section 2.3. In particular, we introduce an asymmetry parameter, $\theta_{4,r}$.[2] In this case, as we show next, exact solutions for $\boldsymbol{\phi}^*$ are achieved taking the log of the generalized logistic function—as stated in equation (O.B.8).

Then, picking a region (here: $\mathcal{T}$) as reference, we apply our normalization procedure. That is, we map the outcome path of the non-reference region $\mathcal{C}$ onto that of the reference region $\mathcal{T}$ using pre-policy data only. To do so, we postulate a normalized path for the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = \left(f_{\mathcal{C}}(.;\boldsymbol{\omega}) \circ y_{\mathcal{C}} \circ t_{\mathcal{C}}(.;\boldsymbol{\psi})\right)(s)$ with

[2] In addition, we drop the level parameter $\theta_{0,r}$ present in the logistic function in Section 2.3.

Figure O.B.4: Decomposition by Normalization Coefficient: Generalized Logistic Paths



Notes: See the notes in Figure 4.

outer composite $f_{\mathcal{C}} = \omega_0 + \omega_1 y_{\mathcal{C}}(t)$ and inner composite $t_{\mathcal{C}} = \psi_0 + \psi_1 s$. That is,

$$\begin{aligned}\widetilde{y}_{\mathcal{C}} &= \omega_0 + \omega_1 y_{\mathcal{C}}\left(\psi_0 + \psi_1 s\right) \\ &= \underbrace{\omega_0 + \omega_1 \ln\left(\theta_{1,\mathcal{C}}\right)}_{=\ln(\theta_{1,\mathcal{T}})} - \underbrace{\omega_1 \frac{1}{\theta_{4,\mathcal{C}}}}_{=\frac{1}{\theta_{4,\mathcal{T}}}} \ln\left(1 + \exp\left(-\underbrace{\theta_{3,\mathcal{C}}\psi_1}_{=\theta_{3,\mathcal{T}}} s + \underbrace{\theta_{2,\mathcal{C}} - \psi_0\theta_{3,\mathcal{C}}}_{=\theta_{2,\mathcal{T}}}\right)\right) = y_{\mathcal{T}}(s)\end{aligned}$$

where the last equality emerges from holding equation (9) with equality, i.e., $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = y_{\mathcal{T}}(s)$ (note that $s = t$ for the reference region $\mathcal{T}$). This implies a system with four equations and four unknowns, $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\} = \{\psi_0, \psi_1, \omega_0, \omega_1\}$:

$$\theta_{1,\mathcal{T}} = \theta_{1,\mathcal{T}}(\omega_0, \omega_1) = \exp\left(\omega_0 + \omega_1 \ln \theta_{1,\mathcal{C}}\right) \tag{O.B.9}$$

$$\theta_{4,\mathcal{T}} = \theta_{4,\mathcal{T}}(\omega_1) = \frac{1}{\omega_1}\theta_{4,\mathcal{C}} \tag{O.B.10}$$

$$\theta_{3,\mathcal{T}} = \theta_{3,\mathcal{T}}(\psi_1) = \psi_1\theta_{3,\mathcal{C}} \tag{O.B.11}$$

$$\theta_{2,\mathcal{T}} = \theta_{2,\mathcal{T}}(\psi_0) = \theta_{2,\mathcal{C}} - \psi_0\theta_{3,\mathcal{C}} \tag{O.B.12}$$

Then, the undetermined normalization coefficients are:

$$\omega_1^* = \frac{\theta_{4,\mathcal{C}}}{\theta_{4,\mathcal{T}}}, \qquad \omega_0^* = \ln\theta_{1,\mathcal{T}} - \frac{\theta_{4,\mathcal{C}}}{\theta_{4,\mathcal{T}}}\ln\theta_{1,\mathcal{C}},$$

$$\psi_1^* = \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}}, \qquad \psi_0^* = \frac{\theta_{2,\mathcal{C}} - \theta_{2,\mathcal{T}}}{\theta_{3,\mathcal{C}}}.$$

That is, the normalization coefficients have an exact and unique solution. For example, let's set the structural parameters $\Theta_{\mathcal{C}} = \{4.00, 5.25, 0.15, 1.00\}$ and $\Theta_{\mathcal{T}} = \{3.50, 7.70, 0.14, 0.85\}$. This implies $\boldsymbol{\phi}^* = \{\psi_0^*, \psi_1^*, \omega_0^*, \omega_1^*\} = \{-16.33, 0.93, -0.37, 1.17\}$. We show the role of each of these normalization coefficients in the non-orthogonal decomposition in Figure O.B.4.

O.C Solution Algorithms for the Economic Models

The Econ-Epi Model. Here we discuss the solution of the model posed in the paper Section 3.1.1.[3] Given t, the amount of economic activity h_z is determined by the following condition,

$$\underbrace{\frac{\partial u(c_z, h_z; \chi)}{\partial c_z} w}_{\substack{\text{Marginal Benefit of Working:} \\ \text{Consumption Gain}}} - \underbrace{\frac{\partial u(c_z, h_z; \chi)}{\partial h_z}}_{\substack{\text{Marginal Cost of Working:} \\ \text{Loss of Leisure}}} = \underbrace{\delta \frac{\partial \phi_{\mathcal{P}}(h_z)}{\partial h_z} u(c_{z+1}, h_{z+1}; \chi)}_{\substack{\text{Marginal Cost of Working:} \\ \text{Loss of Lives}}} \qquad \forall z \geq t, \tag{O.C.1}$$

stating that the marginal benefit of working (more consumption) equates its marginal costs consisting of an intratemporal component (disutility from working) and an intertemporal component (loss of lives). We separately solve for the pre-pandemic equilibrium at $t = 0$ (actually, for any $t \leq 0$) before the unexpected arrival of the pandemic at $t = 1$. In this pre-pandemic era there are no infections and, hence, $\phi_i(h_0) = 1$. That is, the equilibrium labor supply sets the right-hand side of the Euler equation (O.C.1) to zero in which case h_0 simply solves an intra-temporal trade-off. The same equilibrium emerges after the pandemic at some large $t = T$ which delivers a terminal condition $h_T = h_0$. We use the following steps to solve this model: (1) solve for hours worked in the pre-pandemic steady state ($\bar{h}$); (2) select the number of periods to simulate $\mathfrak{T}$ picking a large number T and setting $h_0 = h_T = \bar{h}$; (3) given parameters Θ, guess a sequence $\{h_t\}_{t=0}^{T}$; (4) with $\{h_t\}_{t=0}^{T}$, compute sequences for S_t, I_t, R_t, D_t, N_t; (5) use the above sequences to back out a new sequence for $\{h_t\}_{t=0}^{T}$ using (O.C.1). and solving backwards; (6) if the new sequence $\{h_t\}_{t=0}^{T}$ is different than the guess in Step 3, update the guess and go back to Step 4; (7) store the second value of the sequence, namely h_1, and set $h_0 = h_1$ in order to then go back to Step 3 and repeat this step $\mathfrak{T}$ times constructing a solution sequence $\{h_{sol,t}\}_{t=1}^{\mathfrak{T}}$ using all values stored and simulating the underlying epidemic dynamics associated to the solution sequence; (8) to obtain the series with the effect of policy: set all values of the solution sequence after t_p equal to h_{pol}, and simulate the epidemic dynamics. Further, in the case of endogenous policy, as in the paper Section 3.2.2, the policy hits when the cumulative number of deaths reaches a certain number $\overline{D}$. In Step 5 solve backwards only from the date in which policy is implemented, continue with the rest of the steps as described above. Further, to obtain the series without policy, use as initial values those immediately before the policy hits, go to Step 3 and solve for a new sequence of h assuming the policy constrain is never binding.

The Pill Model. We follow the economic environment in the paper Section 3.1.2. Plugging (15) and (16) into (14), we derive the FOC of h as,

$$FOC(h): \qquad \underbrace{q}_{\text{Marginal Cost of Human Capital}} = \underbrace{(1 - \tau(n)) w z_t e_h(h)}_{\text{Marginal Benefit of Human Capital}}, \tag{O.C.2}$$

where the price of human capital equates the marginal benefit consisting of a wage premium net of the costs of children. Since q is constant, the marginal benefit trades off n and h, i.e., a technology that reduces n enhances human capital. The FOC for sexual intercourse x is:

$$FOC(x): \qquad \underbrace{\tau_n(n)\phi_x(x)(1 - \mathbf{1}_{t \geq t_p} g(o)) w (1 + z_t e(h))}_{\text{Marginal Cost of Intercourse}} = \underbrace{\zeta + \kappa_t \phi_x(x)(1 - \mathbf{1}_{t \geq t_p} g(o))}_{\text{Marginal Benefit of Intercourse}}, \tag{O.C.3}$$

where the marginal benefit considers the additional utility from sex and children. The marginal cost reflects the cost of children in terms of human capital. The FOC for pill use o is:

$$FOC(o): \qquad \underbrace{\tau_n(n)\phi(x)\mathbf{1}_{t \geq t_p} g(o) w (1 + z_t e(h))}_{\text{Marginal Benefit of Pill}} = \underbrace{\kappa_t \phi(x) \mathbf{1}_{t \geq t_p} g_o(o) + \iota}_{\text{Marginal Cost of Pill}} \tag{O.C.4}$$

[3] Note that we pose the planners problem in sequential formulation. In particular, the sequence of beliefs on the odds of infection $\xi_{\mathcal{P}}$ is either constant or deterministic. The presence of a stochastic component, e.g. noisy signals for belief's updates on the odds of infection (Aleman et al., 2022), typically require the problem to be expressed in recursive formulation.

where the marginal cost of the pill is a reduction of utility derived from children and the marginal benefit of the pill is a reduction in the price of human capital.

We use the set of FOC's together with parameters $\Theta = \{\kappa, \zeta, q, w, z, \{\lambda_t\}_{t=1}^T, \{\iota_t\}_{t=1}^T, \{\kappa_t\}_{t=1}^T, \theta_x, \theta_h, \theta_o\}$, to simulate paths for x_t, o_t and h_t and subsequently obtain paths for fertility $n(x_t)$ and share of women with college $e(h_t)$.

Growth Policy and Structural Transformation Model. Following our discussion on structural transformation model in the paper Section 3.1.3, we derive three first order conditions for the household problem.[4] First, an intratemporal condition governing the substitution across consumption goods:

$$FOC(c_{at}): \qquad u_{c_{at}}(c_{at})\frac{1}{p_{at}} = \kappa v_{c_{mt}}(c_{mt}) \tag{O.C.5}$$

Second, an intertemporal Euler condition for k_{t+1} governing the trade off between one additional unit of consumption today versus tomorrow's consumption,

$$FOC(k_{t+1}): \qquad u_{c_a}(c_{at})\frac{1}{p_{at}} = \beta u_{c_a}(c_{at+1})\frac{1}{p_{at+1}}\left(1 + r_{t+1} - \delta\right), \tag{O.C.6}$$

and note that we can rewrite this intertemporal condition in terms of c_m using (O.C.5). Third, an intratemporal condition for n_a equates wages across sectors,

$$FOC(n_{at}): \qquad u_{c_a}(c_{at})(w_{at} - w_{mt}) = 0 \tag{O.C.7}$$

These allocations need to satisfy the marginal product conditions arising from the firms' problems in competitive markets, that is, $w_{at} = \phi\frac{p_{at}y_{at}}{n_{at}}$, $w_{mt} = \alpha\frac{y_{mt}}{n_{mt}}$ and $r_t = (1-\alpha)\frac{y_{mt}}{k_t}$.

We solve this model by guessing the sequences of factor prices $\{w, r\}_{t=0}^{\infty}$ with $w_t = w_{at} = w_{mt}$. Given these prices, we find the allocations c_{at}, k_{t+1} and n_{at} that solve the set of first order conditions (O.C.5)-(O.C.7) with $p_{at} = \frac{w_t}{\phi z_{at}}\left(\frac{n_{at}}{\ell}\right)^{1-\phi}$.[5] There is market clearing in labor and capital, and aggregate consistency. Note that the intertemporal Euler condition (O.C.6) is a second order difference equation in $\{k_t, k_{t+1}, k_{t+2}\}$ at every period t. We use as initial and terminal conditions the corresponding stationary economies at $t = 0$ and at a large T with negligible agricultural share of labor.

O.D Robustness to the Availability of Pre-Policy Data

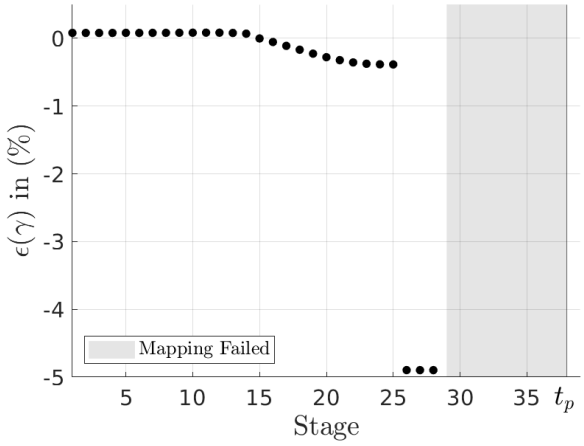
In our models and applications, we have not encountered situations in which the availability or lack of pre-policy data prevented our normalization procedure from generating a successful mapping. In this Appendix, we put this to the test, particularly because of Roth (2022). Precisely, we conduct an experiment in which we explicitly reduce the number of pre-policy data points available to the policy evaluator. This allow us to assess how the error in which SBI captures the model-generated policy effects changes with the amount of pre-policy data.

For the experiment, we used the benchmark econ-epi model of Section 3.1.1. In this benchmark, there are $t_p - t_0$ periods available before policy is implemented in the reference region, with $t_0 = 0$ and $t_p = 38$. We artificially reduce the number of periods available to the policy evaluator to be $t_p - (t_0 + n)$ for $n \in [0, t_p - 1]$. We then re-conducted SBI for each n, showing the result in Figure O.D.1. This plot shows how the policy error changes with the amount of pre-policy data.

[4] Note that we can isolate c_m from (18) and plug it into (17) plus use $n_{mt} = 1 - n_{at}$. This implies that we can maximize the objective function in terms of the sequences of three unknowns $\{c_{at}, n_{at}, k_{t+1}\}_{t=0}^{\infty}$.

[5] Note that without the distortion τ, if $\phi = \alpha$ and if we had the same factor inputs in the production of both goods, then the equality of the ratio of factor input prices across sectors would imply that the ratio of factor inputs must be identical across sectors. In turn, this would imply a standard result for the pricing of agricultural goods, $p_{at} = \frac{z_{mt}}{z_{at}}$, which renders the price of agricultural good as exogenous. The fact that we allow for ϕ to differ from α and that we have different factor inputs differ across sectors both prevents the standard result. Indeed, in our case, the price of agricultural goods depends endogenously on n_{at}.

Figure O.D.1: SBI Performance with Limited Data Available Before Policy Implementation



Notes: Vertical axis: Policy error as constructed in Section 2. Horizontal axis: Stage (time in the reference region, here: $\mathcal{T}$) in which the data (in all regions) starts to be available for the policy evaluator. The shaded gray area denotes the data availability for which the normalization ceased to work.

We find that the normalization procedure works well in our econ-epi model, and the policy error remained below 0.5% until $n = 25$, and below 5.0% until $n = 28$. However, the normalization procedure ceased to work at $n = 29$. This means that, in our calibrated model, if the policy evaluator only has $t_p - n = 38 - 29 = 9$ data points before policy, then the time series of pre-policy data is too short to successfully conduct the normalization procedure and pin down the policy effects with SBI. Here, we would like to emphasize that the minimum amount of data points before policy implementation (10 periods in our illustration) that are needed for SBI to be able to recover the true policy effects is likely to depend on the shape of the outcome time paths. This analysis is something we plan to explore in future work.[6]

O.E Analysis of Confounding Factors

Here, we discuss the scenarios which we use to assess how SBI performs in the presence of confounding factors, referenced to in Section 3.2.2 of the paper. First, in Appendix O.E.1, we show the path of time-varying latent heterogeneity that we assumed for our analysis in the main text. Second, in Appendix O.E.2, we assess the implications of confounding policy for the performance of SBI.

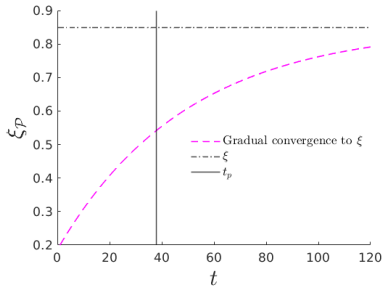
O.E.1 Time-Varying Latent Heterogeneity

Consider a scenario in which one region exogenously, $\mathcal{T}$, learns about the process of infection throughout the epidemic. That is, we allow for the beliefs on the infection process, $\xi_{\mathcal{P}}$, to exogenously and gradually approach the actual ξ in region $\mathcal{T}$ but not in region $\mathcal{C}$. In Figure O.E.1, we show the assumed time path for $\xi_{\mathcal{P}}$ (magenta dashed line) converging from below (from $t = 0 \leq t_p$) to the true probability ξ (dashed gray line). Hence, a structural parameter, $\xi_{\mathcal{P}}$, that is unobserved to the policy evaluator evolves over time before and after policy implementation.

Due to learning, there is a larger behavioral response (reduction of hours) to the pandemic in region $\mathcal{T}$ relative to the scenario with a fixed $\xi_{\mathcal{P}}$, see panel (a1) in Figure O.E.2. We show the implied true policy effects on the flow

[6] Additionally, we note that SBI may potentially work with fewer data points if they are available at a lower frequency but still able to capture changes in curvature, mainly, the increasing and decreasing growth of the outcome path before policy.

Figure O.E.1: Time-Varying Latent Heterogeneity: Assumed Path of Beliefs $\xi_{\mathcal{P}}$ for region $\mathcal{T}$



Notes: Figure shows the time path of beliefs $\xi_{\mathcal{P}}$.

of deaths in panel (a2), the identified policy effects in panel (a3) and a comparison between true and identified effects in panel (a4) of Figure O.E.2. The identified effects imply $\gamma = 12.242\%$ of lives saved which is close to the true effects, $\gamma_{\text{true}} = 12.663\%$ i.e., a policy error of 3.318%;. Thus, SBI can recover the policy effects in contexts where there is time-varying latent heterogeneity. At the same time, analogously to the discussion in Section 3.2.1, the robustness of our method to time-varying latent heterogeneity is bounded by how far this heterogeneity drives the regional outcome paths away from each other.[7]

O.E.2 Additional Confounding Policy

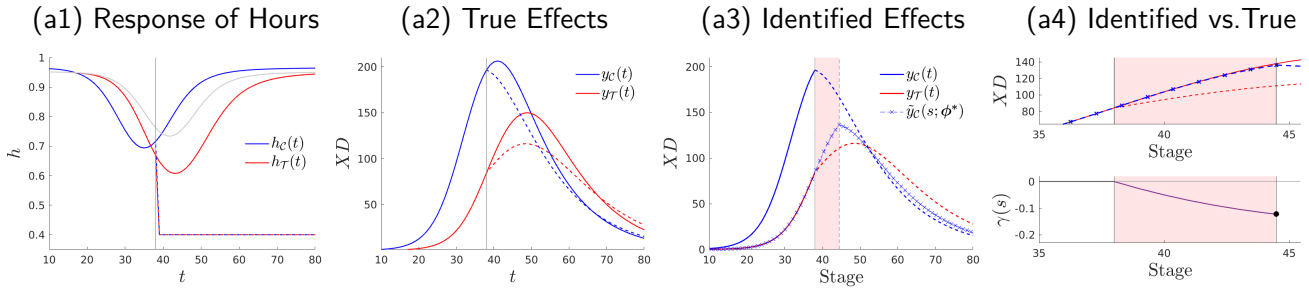
Consider a scenario in which we introduce an unanticipated policy that imposes an additional (and weaker) constraint on hours worked, $\bar{h} = 0.71$, in region $\mathcal{T}$ one period before the nationwide policy under evaluation (i.e., $\bar{h} = 0.40$ in all regions at t_p) is put in place; panel (b1), Figure O.E.2. We apply SBI to assess the nationwide policy introduced at t_p, while purposefully ignoring the presence of the confounding policy; panel (b2). The identified effects are 14.708% of lives saved which is close to the true effects in that window, 15.588%; i.e., a policy error of 5.644%; panels (b3) and (b4). Thus, SBI can recover the true policy effects in a context where there exists a confounding policy.

Alternatively, consider the opposite scenario in which confounding policy is implemented in region $\mathcal{C}$ (and not in region $\mathcal{T}$) before the stay-home policy under study takes place. In this case, we find that the policy effect is recovered with an error of 2.34%; see Figure O.E.3. Further, for both scenarios, in line with Section 3.2.1, it is straightforward to show that the policy error increases if the effect of the confounding policy forces the cross-regional outcome paths to be sufficiently dissimilar before t_p.

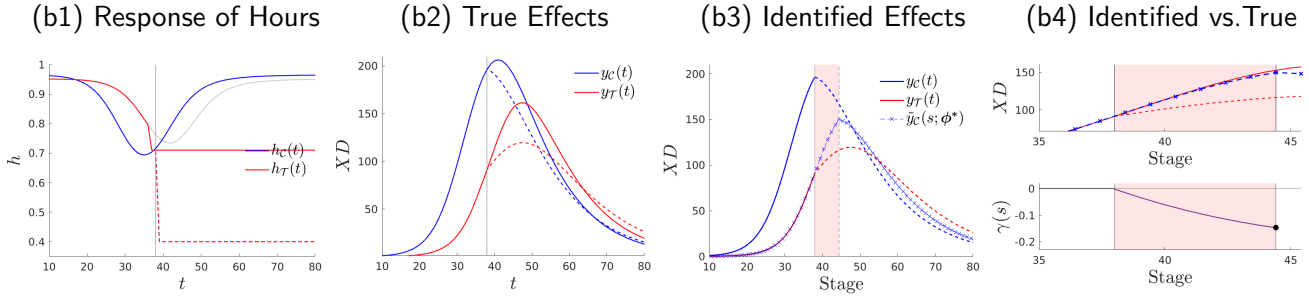
[7] See Online Appendix O.F for the discussion of the bounds under which outcome divergence after policy due to structural parameters (and not to policy) alters the ability of SBI to capture the true policy effects, as the time path of $\xi_{\mathcal{P}}$ is a potential source of structural divergence, as in Rambachan and Roth (2023).

Figure O.E.2: Stage-Based Identification of Policy Effects: Confounding Factors

(a) Time-Varying Latent Heterogeneity:

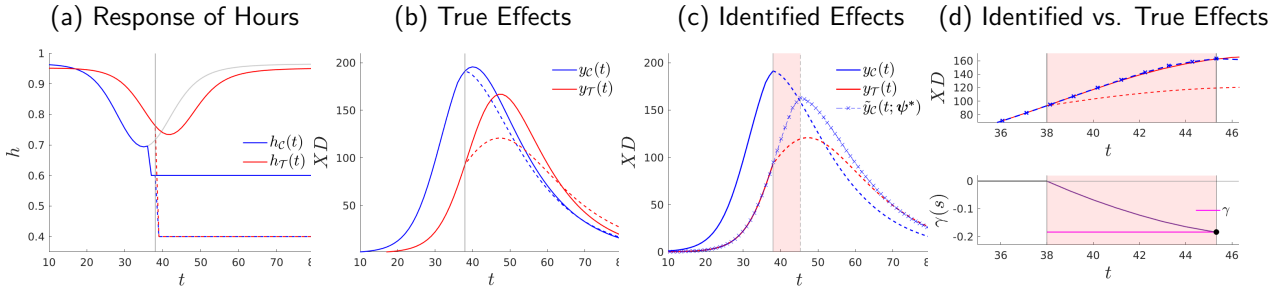


(b) Confounding Policy (in $\mathcal{T}$):



Notes: Panel (a) uses a path for time-varying latent heterogeneity in region $\mathcal{T}$ as described in the text. Panel (b) has a confounding policy of $\bar{h} = 0.71$ at $t = t_p - 1$ in $\mathcal{T}$.

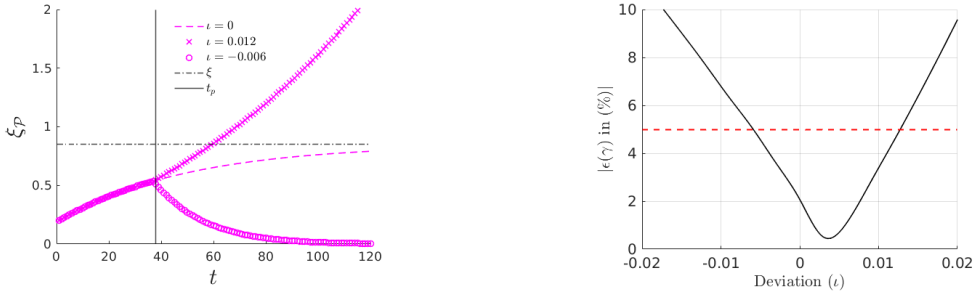
Figure O.E.3: Stage-Based Identification of Policy Effects: With Confounding Policy in $\mathcal{C}$



Notes: Shows outcome paths, where $\bar{h}_1 = 0.6$ at $t = t_p - 1$ in $\mathcal{C}$ (this policy is unobserved), and $\bar{h}_2 = 0.4$ (lockdown) at t_p, $\gamma = -18.34\%, \epsilon(\gamma) = 2.34\%$.

Figure O.F.1: Structural Divergence of Outcome Time Paths: Pre- Vs. Post-Policy

(a) Time-Varying Latent Heterogeneity, $\xi_{\mathcal{P}}$ (b) Policy Error and Structural Divergence



Notes: Panel (a) shows the evolution of beliefs for different values of ι. Panel (b) shows the error of the estimated policy effect as a function of the value of ι for region $\mathcal{T}$, while holding $\iota = 0$ in region $\mathcal{C}$.

O.F Structural Divergence of Outcome Paths

An important concern for the credibility of empirical strategies is that the patterns observed after a policy change may not be solely due to the policy itself but to the fact that the regional outcomes diverge structurally over time, even in the absence of policy. This was carefully studied in Rambachan and Roth (2023). In the context of our research on SBI, we addressed this issue in part by evaluating the performance of our method when there is time-varying latent heterogeneity, in which structural parameters potentially differ across regions and time, as discussed in Section 3.2.2. In this Appendix, we extend that analysis.

Building on the econ-epi model of Section 3.1.1, we assume that there is structural time-varying latent heterogeneity across regions in the beliefs ($\xi_{\mathcal{P}}$) about the actual odds of infection (ξ). Specifically, we extend our previous analysis by assuming that the path for beliefs in one region (here: region $\mathcal{T}$) evolves according to $\xi_{\mathcal{P},t+1} = (1+\iota)^{t-t_p}\xi_{\mathcal{P},t}$ if $t \geq t_p$, where ι captures the belief's growth rate, a source of structural divergence across regions, for periods after policy. The beliefs' time paths for region $\mathcal{T}$ are shown in panel (a) of Figure O.F.1, where we assume that policy is not driving the behavior of the beliefs and set $\iota = 0$ for region $\mathcal{C}$. Our aim is to determine the values of ι that enable SBI to capture the true (model-generated) policy effect with an error $|\epsilon(\gamma)| < 5\%$. As discussed in Section 3.2.2, we know that SBI recovers the true policy effects as long as the underlying time-varying heterogeneity, in this case $\xi_{\mathcal{P}}$, does not cause the regional outcome paths to diverge too much in the absence of policy. Thus, the policy error will be lower if the values of ι are not too different from zero. Consequently, ι constrains SBI's ability to recover the true policy effects. The results of this illustration are presented in panel (b) of Figure O.F.1. As long as the values of ι are within the range of $[-0.006, 0.012]$, SBI recovers the true policy effects with an error below 5%.

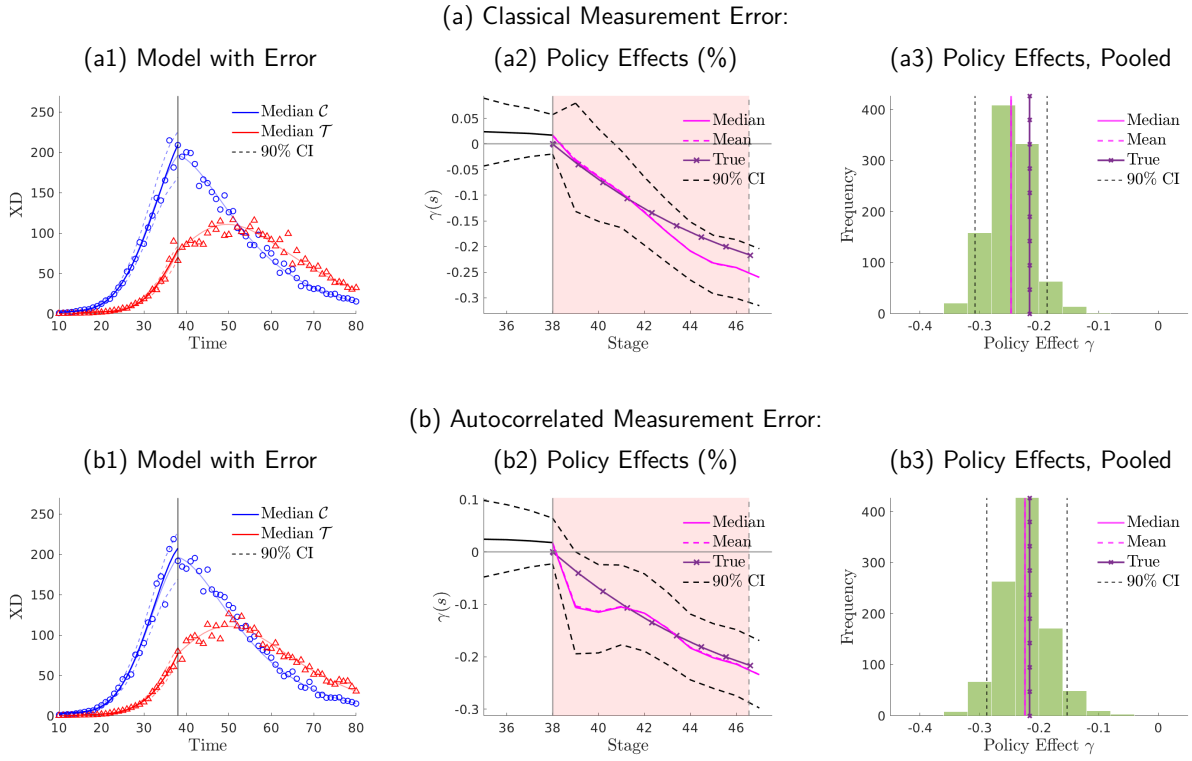
Finally, it is important to note that the normalization procedure is conducted using only data from before policy implementation ($t < t_p$). As a result, the normalization is identical across different values of ι in our illustration. Therefore, even if the normalized outcome paths are not significantly different from each other in the stage dimension before policy implementation, this does not guarantee success in recovering the true policy effects. This is an analogous result to that highlighted earlier in Rambachan and Roth (2023) for standard difference-in-difference frameworks, which we believe deserves further exploration in the context of SBI settings. We leave this as a topic for future research.

O.G Alternative Inference Procedure

In Section 3.4, we discuss how we conduct inference in our applications throughout our applications. In this Appendix, we discuss an altogether alternative way to conduct inference. This alternative way proceeds by using the recovered estimates for the error terms $u_r(t)$ in order to estimate the sample variance of the errors, i.e.,

$\widehat{\sigma}_r$. Then, under a normality assumption on the error term in (19), we simulate $Q = 1,000$ paths of errors and, hence, the same number of pre-policy outcome paths onto which we apply the smoothing step in order to recover a simulation-specific estimand $\widehat{\widehat{y}}_{r,q}(t)$. Since the estimand $\widehat{\widehat{y}}_{r,q}(t)$ differs by simulation $q \in Q$, each simulation delivers a stage-based identified policy effect, γ_q. We show the results of this different inference in Figure O.G.1. Overall, we find similar insights with an identified mean policy effect of 25.20% [19.62,30.91] and 22.60% [15.91,28.73] with, respectively, classical measurement error and with auto-correlated measurement error. The recovered policy effect is not significantly different from the true (model-generated) policy effect.

Figure O.G.1: Stage-Based Identification of Model-Generated Policy Effects: Alternative Inference



Notes: We use the benchmark calibration in the paper Section 3.1.1. the top panels (a), we introduce classical measurement error in our model with $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$. In the bottom panels (b), we introduce non-classical measurement error with $\{\rho_{\mathcal{C}}, \rho_{\mathcal{T}}\} = \{0.13, 0.13\}$ and $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$.

O.H Robustness to the Trend-Extraction

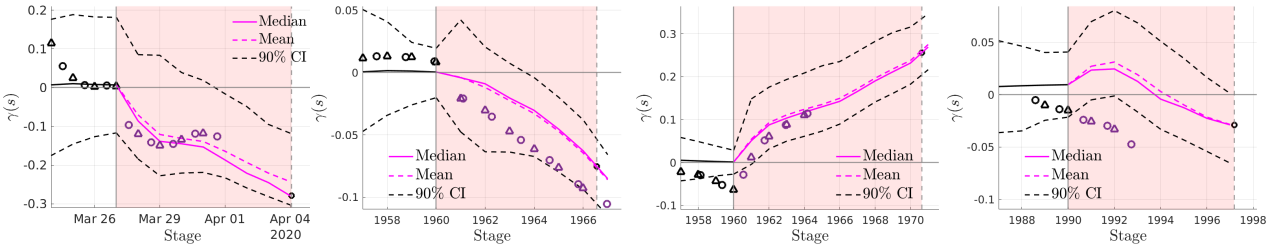
We use the trend-extraction step described in Section 4 in order to conduct inference using as stochastic component the empirical deviations from trend. In Figure O.H.1, we compare the policy effects of each empirical application in which we used a trend-extraction step with a counterpart where we skip the trend-extraction step and minimized the distance between original data points. We find that the trend extraction does not affect neither the identified effects of the Spanish Confinamiento (see panel (a) of Figure O.H.1) nor the effects of the 1960 FDA approval of oral contraceptives (see panel (b) and (c) of Figure O.H.1). However, we find that the identified policy effects without the trend-extraction step differ from those with the trend-extraction in the context of the German reunification; see panel (d) in Figure O.H.1. The reason for this differential result is that in the case of the German reunification the outcome variable (GDP per capita) shows fluctuations at a larger frequency (business

cycles) than the frequency in which we are ultimately interested in for the evaluation of the German reunification. Since all regions share the same aggregate fluctuations, if we do not purge GDP per capita from the business cycle fluctuations the normalization algorithm that aims to minimizes the distance between the GDP per capita across regions (to some reference region) will be drawn to map the business cycle fluctuations of all regions to that of the reference region. Therefore, in addition to be useful to conduct inference, the trend-extraction step can also serve the purpose of removing fluctuations that are of higher frequency than the ones in which the policy evaluator is interested in.

Figure O.H.1: Robustness to Trend-Extraction

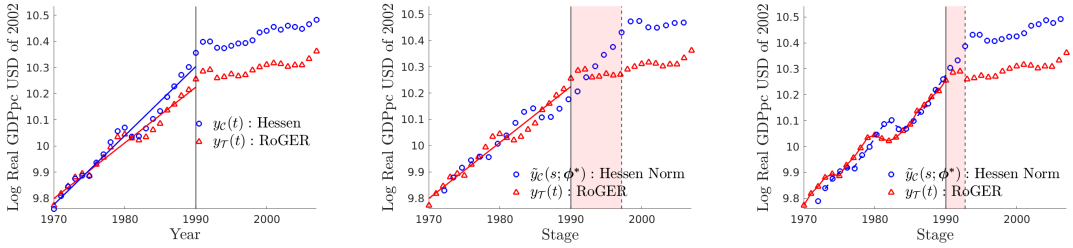
(a) Identified Policy Effects Without the Trend-Extraction Step:

(a1) Spanish *Confinamiento* (a2) Crude Birth Rate (a3) Women College Attainment (a4) German reunification



(b) Trend Extraction and the Effects of the German reunification:

(b1) Before Normalization (b2) After Normalization (w/ trend extraction) (b3) After Normalization (without trend extraction)



Notes: We show the policy effects identified without the implementation of the smoothing step and around the neighborhood of the median estimate, as described in Section 4.

O.I Many Control Regions Per Treated Region

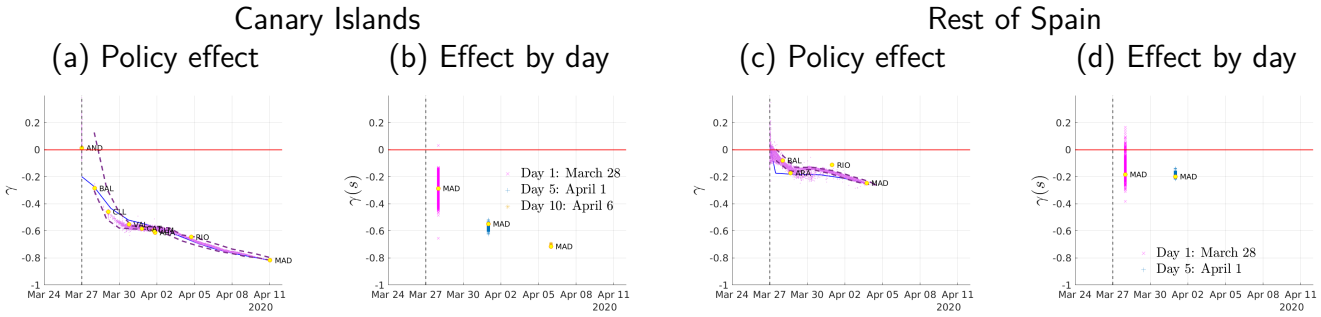
In Section, 5.1, we assess heterogeneity of policy effects following a one-control-many-treatments approach in which we mapped the pre-policy time paths of all regions (and hybrids) onto that of the stage-leading region (Madrid). Clearly, by alternatively mapping the pre-policy time path of the stage-leading region separately onto that of each lagging region, we can also identify heterogeneity of policy effects by region. For each lagging regions, however, multiple candidate control regions exist: all regions that lead it in terms of stages. In this Section we show that the identified policy effect is robust to the choice of the control region.

Consider the region of Canary Islands, which received the Spanish national lockdown at a relatively early stage of the Covid-19 wave. We choose it as reference region and (one-by-one) map a large set of non-reference regions onto the outcome path of the Canary Islands. Again, the set of potential controls is composed by all hybrid regions that are part of the power set of Spanish regions without Canary Islands (minus the empty set), i.e., $2^{16} - 1 = 65,535$ hybrid regions. From this set we only keep those which, after the normalization step, turn out to lead the Canary Islands in terms of stages which results in a total of 64,903 controls for the Canary Island. In panel (a) of Figure 8, we show the identified overall policy effect of the stay-home policy on the flow of deaths

in the Canary Islands that results from using control regions that at the effective policy date $t_p + \tau$ are at the stage indicated on the x-axis. Recall that the stage now is "Canary Islands time". The solid blue line shows the full trajectory of $\gamma(s)$ using Madrid as a control, with the last point corresponding to the point estimate reported in panel (b) of Figure 16. We repeat the same exercise for the hybrid region RoSPA ($2^{17} - 1 = 131,071$), and show the results in panel (c) of Figure 8.[8] For the specific case of RoSPA we find that the identified policy effect (cf. Section 4.1) falls within the 95% CI [-25.70,-23.88] emerging from the distribution of policy effects obtained using multiple (all potential) controls for RoSPA.[9]

The crucial take-away from panels (a) and (c) of Figure 8 is that for a given length of the identification window, which control region to use as counterfactual has little influence: the variance of the estimated $\gamma(s)$ is very small. Note that the identified policy effect is larger when using a control that is more advanced. Panels (b) and (d) show the corresponding effects by day. As done in section 5.1 one can keep track of the effects by stage estimated using a particular region as control. The yellow dots in panels (b) and (d) keep track of $\gamma(s)$ using Madrid as a control; note that the values in panel (d) are the same as their counterpart in Figure 16 when mapping Madrid onto RoSPA.

Figure O.I.1: The Effects of the Spanish *Confinamiento* in the Canary Islands and Rest of Spain, with Different Controls



Notes: In panels (a) and (b) the reference is Canary Islands and in panels (c) and (d) Rest of Spain. Each point represents the policy effect for Canary Islands or Rest of Spain respectively as estimated with a control region that received policy on the indicated stage. The dashed lines are 90% confidence bands at the daily frequency.

O.J Non-Nationwide Policy: Model Analysis

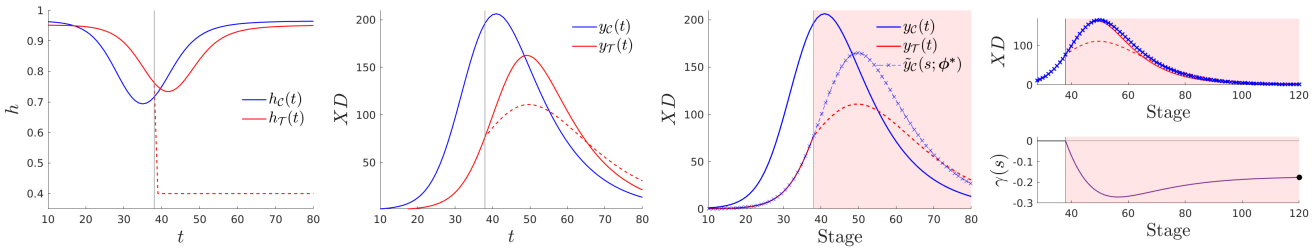
As an illustration, we modify the benchmark experiment of Section 3.1.1 introducing the stay-home policy in region $\mathcal{T}$, but not in region $\mathcal{C}$; see the implications for hours and deaths in, respectively, panel (a) and (b) of Figure O.J.1. The results of the normalization are shown in panel (c). In the example, we find that region $\mathcal{C}$ leads region $\mathcal{T}$ at the time of policy implementation. The identified policy effects overlap with the true (model-generated) effects; see panel (d), Figure O.J.1.

[8] Notice from panel (c) of Figure that the ranking of regions that lead RoSPA is different from the ranking displayed in panel (c) of Figure 16. E.g., here the Baleares (BAL) is leading RoSPA whereas in Figure 16 it is lagging it. This results from the fact that the set of normalization coefficients is overidentified, and the choice of the reference region can affect the part of the regional time paths that end up in interval $\mathbb{C}(s)$. Thus, while transitivity of rankings of regions holds across choices of reference regions in scenarios with exact identification, it does not apply generally. However, for regions "sufficiently" far in stages from RoSPA at the time of policy implementation transitivity continues to hold and therefore the general message of Figure 8 is not affected by these violations.

[9] To obtain the entire set of cross-regional heterogeneous policy effects one can repeat the exercise by region for every non-leading region.

Figure O.J.1: Stage-Based Identification of Model-Generated Effects: A Non-Nationwide Policy

(a) Response of Hours (b) True Effects (c) Identified Effects (d) Identified vs.True



Notes: Where $\bar{h} = 0.4, t_p = 38, t_f = 250$ with $\gamma = 17.663\%$ and $\epsilon(\gamma) = 5.819\%$

O.K Alternative Lemma for Proof of Theorem 1

Here, we discuss an additional Lemma, which complements Lemma 1 and can be seen as an alternative to Lemma 2 in the proof of Theorem 1 presented in Appendix A. In the case of $\omega_0^* \neq 0$, we can establish equivalence of the identified policy effects for a slightly modified policy effect. This modified effect subtracts from each path its value at the policy date t_p expressed in region $\mathcal{T}$-time. Then, we express the policy effect based on those modified paths. When choosing $\mathcal{T}$ as the reference region, the resulting paths are $(y_{\mathcal{T}}(s) - y_{\mathcal{T}}(s = t_p))$ and $(\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - \widetilde{y}_{\mathcal{C}}(s = t_p;\boldsymbol{\phi}^*))$, where again $y_{\mathcal{T}}(s = t_p) = \widetilde{y}_{\mathcal{C}}(s = t_p;\boldsymbol{\phi}^*) = \omega_0^* + \omega_1^* y_{\mathcal{C}}(t_{\mathcal{C}}(s = t_p, \boldsymbol{\psi}^*))$ up to a minimization error. This gives the modified policy effect

$$\begin{aligned}\gamma^{mod} &= \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \left((y_{\mathcal{T}}(s) - y_{\mathcal{T}}(t_p)) - (\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - \widetilde{y}_{\mathcal{C}}(t_p;\boldsymbol{\phi}^*))\right) ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - \widetilde{y}_{\mathcal{C}}(t_p;\boldsymbol{\phi}^*)) ds} \\ &= \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*))\, ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - y_{\mathcal{T}}(t_p)) ds}\end{aligned} \tag{O.K.1}$$

When choosing $\mathcal{C}$ as the reference region, the resulting paths are $\left(y_{\mathcal{C}}(\boldsymbol{s}) - y_{\mathcal{C}}(\boldsymbol{s} = t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*))\right)$ and $\left(\widetilde{y}_{\mathcal{T}}(\boldsymbol{s};\bar{\boldsymbol{\phi}}^*) - \widetilde{y}_{\mathcal{T}}(\boldsymbol{s} = t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*))\right)$. Note in the previous two expressions that the paths are subtracting the value at the policy date in region $\mathcal{T}$-time. Using those in (O.K.1) (adjusted to region $\mathcal{C}$ as reference region) again gives rise to the same (modified) policy effect which we summarize in the following

Lemma 3. *If $\omega_0^\star \neq 0$, then the choice of the reference region is irrelevant for the policy effect when expressing it in adjusted form* (O.K.1)*, i.e.,*

$$\gamma^{mod} = \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*))\, ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - y_{\mathcal{T}}(t_p)) ds} = \frac{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} \left(\widetilde{y}_{\mathcal{T}}(\boldsymbol{s};\bar{\boldsymbol{\phi}}^*) - y_{\mathcal{C}}(\boldsymbol{s})\right) d\boldsymbol{s}}{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} (y_{\mathcal{C}}(\boldsymbol{s}) - \widetilde{y}_{\mathcal{T}}(\boldsymbol{s} = t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*)) d\boldsymbol{s}}. \tag{O.K.2}$$

Proof. The proof follows the exact same steps as before. In addition, we need to consider the modified denominator. For this, note that

$$\widetilde{y}_{\mathcal{T}}(\boldsymbol{s} = t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*)) = \bar{\omega}_0^* + \bar{\omega}_1^* y_{\mathcal{T}}\left(t_{\mathcal{T}}(t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*);\bar{\boldsymbol{\psi}}^*)\right) = \bar{\omega}_0^* + \bar{\omega}_1^* y_{\mathcal{T}}(t_p) \tag{O.K.3}$$

Thus, plugging this into the RHS expression of (O.K.2) we obtain

$$\frac{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} \left(\widetilde{y}_{\mathcal{T}}(\boldsymbol{s};\bar{\boldsymbol{\phi}}^*) - y_{\mathcal{C}}(\boldsymbol{s})\right) d\boldsymbol{s}}{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} y_{\mathcal{C}}(\boldsymbol{s}) - \widetilde{y}_{\mathcal{T}}(\boldsymbol{s} = t_{\mathcal{T}}^{-1}(t_p;\bar{\boldsymbol{\psi}}^*))d\boldsymbol{s}} = \frac{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} \left(\bar{\omega}_1^* y_{\mathcal{T}}(t_{\mathcal{T}}(\boldsymbol{s},\bar{\boldsymbol{\psi}}^*)) - \bar{\omega}_1^* \widetilde{y}_{\mathcal{C}}(t_{\mathcal{T}}(\boldsymbol{s},\bar{\boldsymbol{\psi}}^*),\boldsymbol{\phi}^*)\right) d\boldsymbol{s}}{\int_{\mathbb{W}(\boldsymbol{s};\bar{\boldsymbol{\psi}}^*)} (\bar{\omega}_0^* + \bar{\omega}_1^* \widetilde{y}_{\mathcal{C}}(t_{\mathcal{T}}(\boldsymbol{s},\bar{\boldsymbol{\psi}}^*),\boldsymbol{\phi}^*)) - (\bar{\omega}_0^* + \bar{\omega}_1^* y_{\mathcal{T}}(t_p))d\boldsymbol{s}}$$
$$= \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \left(y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s,\boldsymbol{\phi}^*)\right) ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s,\boldsymbol{\phi}^*) - y_{\mathcal{T}}(t_p) ds}.$$

□

References

Aleman, C., Iorio, D., and Santaeulàlia, R. (2022). A Quantitative Theory of the HIV Epidemic: Education, Risky Sex and Asymmetric Learning. Working paper, Barcelona School of Economics.

Rambachan, A. and Roth, J. (2023). A More Credible Approach to Parallel Trends. *The Review of Economic Studies*.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–22.