Online Appendix to:

Conditional relationships in dynamic models

Zach Warner*

March 10, 2019

*Postdoctoral Research Fellow, Cardiff University. email: WarnerZ@cardiff.ac.uk, web: http://www.zachwarner.net.

Contents

1 Proofs

1.1 Proposition 1

Proof. Recall that for the model of interest there are only two covariates, x and z, and a response y, each of which is a covariance-stationary AR(1) process. These time series can then be represented individually as

$$y_t = \zeta_y + \rho_y y_{t-1} + \eta_t, \qquad \text{where } |\rho_y| < 1, \tag{1}$$

$$x_t = \zeta_x + \rho_x x_{t-1} + \eta_t, \qquad \text{where } |\rho_x| < 1, \tag{2}$$

$$z_t = \zeta_z + \rho_z z_{t-1} + \eta_t, \qquad \text{where } |\rho_z| < 1, \tag{3}$$

with t indexing time. Assume that $\eta \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$. Following Davidson and MacKinnon (2003) and Box-Steffensmeier et al. (2014), the unconditional mean and variance, and unconditional autocovariance for any lag length ℓ, can be written

$$\mathbb{E}[y_t] = \frac{\zeta_y}{1-\rho_y}, \qquad \mathbb{V}[y_t] = \frac{\sigma^2}{1-\rho_y^2}, \qquad \mathbb{C}[y_t, y_{t+\ell}] = \rho_y^\ell \left(\frac{\sigma^2}{1-\rho_y^2}\right), \tag{4}$$

$$\mathbb{E}[x_t] = \frac{\zeta_x}{1-\rho_x}, \qquad \mathbb{V}[x_t] = \frac{\sigma^2}{1-\rho_x^2}, \qquad \mathbb{C}[x_t, x_{t+\ell}] = \rho_x^\ell \left(\frac{\sigma^2}{1-\rho_x^2}\right), \tag{5}$$

$$\mathbb{E}[z_t] = \frac{\zeta_z}{1-\rho_z}, \qquad \mathbb{V}[z_t] = \frac{\sigma^2}{1-\rho_z^2}, \qquad \mathbb{C}[z_t, z_{t+\ell}] = \rho_z^\ell \left(\frac{\sigma^2}{1-\rho_z^2}\right). \tag{6}$$

Define $\mathcal{F}(\mu, \sigma^2)$ as generic distribution with mean μ, variance σ^2, and autocovariance $\gamma(\ell)$ for lag ℓ. Thus, we can define the variables as stochastic processes using the Equations 4-6 as

$$y \sim \mathcal{F}(\mu_y, \sigma_y^2), \tag{7}$$

$$x \sim \mathcal{F}(\mu_x, \sigma_x^2), \tag{8}$$

$$z \sim \mathcal{F}(\mu_z, \sigma_z^2). \tag{9}$$

I now calculate the unconditional mean, variance, and autocovariance of $x_t z_t$ in turn.

Mean. The expectation is given by

$$\begin{aligned}\mathbb{E}\left[x_t z_t\right] &= \mathbb{C}[x_t, z_t] + \mathbb{E}[x_t]\mathbb{E}[z_t] \\ &= \mathbb{C}[x_t, z_t] + \mu_x \mu_z. \end{aligned} \tag{10}$$

Variance. The variance is given by

$$\begin{aligned}\mathbb{V}[x_t z_t] = &\left(\mathbb{E}[x_t]\right)^2 \mathbb{V}[z_t] + \left(\mathbb{E}[z_t]\right)^2 \mathbb{V}[x_t] + \mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)^2 \left(z_t - \mathbb{E}[z_t]\right)^2\right] \\ &+ 2\mathbb{E}[x_t]\mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)\left(z_t - \mathbb{E}[z_t]\right)^2\right] \\ &+ 2\mathbb{E}[z_t]\mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)^2\left(z_t - \mathbb{E}[z_t]\right)\right] \\ &+ 2\mathbb{E}[x_t]\mathbb{E}[z_t]\mathbb{C}[x_t, z_t] - \left(\mathbb{C}[x_t, z_t]\right)^2,\end{aligned}$$

which reduces to

$$\begin{aligned}\mathbb{V}[x_t z_t] = &\ \mathbb{C}[x_t^2, z_t^2] + 2\mu_x \mathbb{C}[x_t, z_t^2] + 2\mu_z \mathbb{C}[x_t^2, z_t] - \left(\mathbb{C}[x_t, z_t]\right)^2 \\ &+ \left(\sigma_x^2 + 2\mu_x^2 - 2\mu_x\right)\left(\sigma_z^2 + 2\mu_z^2 - 2\mu_z\right). \end{aligned} \tag{11}$$

Auto-covariance. Bohrnstedt and Goldberger (1969) build on results from Goodman (1960), giving the general form of the covariance of two products of random variables:

$$\begin{aligned}&\mathbb{C}[x_t z_t, x_{t+\ell} z_{t+\ell}] = \\ &\quad \mathbb{E}\left[x_t\right]\mathbb{E}\left[x_{t+\ell}\right]\mathbb{C}\left[z_t, z_{t+\ell}\right] + \mathbb{E}\left[x_t\right]\mathbb{E}\left[z_{t+\ell}\right]\mathbb{C}\left[x_{t+\ell}, z_t\right] \\ &\quad + \mathbb{E}\left[z_t\right]\mathbb{E}\left[x_{t+\ell}\right]\mathbb{C}\left[x_t, z_{t+\ell}\right] + \mathbb{E}\left[z_t\right]\mathbb{E}\left[z_{t+\ell}\right]\mathbb{C}\left[x_t, x_{t+\ell}\right] \\ &\quad + \mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)\left(z_t - \mathbb{E}[z_t]\right)\left(x_{t+\ell} - \mathbb{E}[x_{t+\ell}]\right)\left(z_{t+\ell} - \mathbb{E}[z_{t+\ell}]\right)\right] \\ &\quad + \mathbb{E}[x_t]\mathbb{E}\left[\left(z_t - \mathbb{E}[z_t]\right)\left(x_{t+\ell} - \mathbb{E}[x_{t+\ell}]\right)\left(z_{t+\ell} - \mathbb{E}[z_{t+\ell}]\right)\right] \\ &\quad + \mathbb{E}[z_t]\mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)\left(x_{t+\ell} - \mathbb{E}[x_{t+\ell}]\right)\left(z_{t+\ell} - \mathbb{E}[z_{t+\ell}]\right)\right] \\ &\quad + \mathbb{E}[x_{t+\ell}]\mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)\left(z_t - \mathbb{E}[z_t]\right)\left(z_{t+\ell} - \mathbb{E}[z_{t+\ell}]\right)\right] \\ &\quad + \mathbb{E}[z_{t+\ell}]\mathbb{E}\left[\left(x_t - \mathbb{E}[x_t]\right)\left(z_t - \mathbb{E}[z_t]\right)\left(x_{t+\ell} - \mathbb{E}[x_{t+\ell}]\right)\right] \\ &\quad - \mathbb{C}\left[x_t, z_t\right]\mathbb{C}\left[x_{t+\ell}, z_{t+\ell}\right]. \end{aligned} \tag{12}$$

This expression simplifies to

$$\begin{aligned}\mathbb{C}[x_t z_t, &x_{t+\ell} z_{t+\ell}] = \\ &4\mu_x^2\mu_z^2 + \mu_x^2\rho_z^\ell\sigma_z^2 + \mu_z^2\rho_x^\ell\sigma_x^2 + \rho_x^\ell\rho_z^\ell\sigma_x^2\sigma_z^2 \\ &+\mu_x\Big(\mathbb{C}[x_t, z_t z_{t+\ell}] - \mathbb{C}[x_{t+\ell}, z_t z_{t+\ell}]\Big) + \mu_z\Big(\mathbb{C}[x_t x_{t+\ell}, z_t] - \mathbb{C}[x_t x_{t+\ell}, z_{t+\ell}]\Big) \\ &-\mu_x\mu_z\Big(\mathbb{C}[x_t, z_t] + \mathbb{C}[x_{t+\ell}, z_{t+\ell}]\Big) + \mathbb{C}[x_t x_{t+\ell}, z_t z_{t+\ell}] + \mathbb{C}[x_t, z_t]\mathbb{C}[x_{t+\ell}, z_{t+\ell}].\end{aligned} \tag{13}$$

Checking for stationarity. Recall that the definition of weak (covariance) stationarity is that, for a stochastic series, the unconditional mean and variance, and unconditional autocovariance for some lag ℓ, must exist and be independent of t (for $t, \ell \in \mathbb{Z}$).

The stochastic process xz satisfies these conditions only where the covariance terms in Equations 10, 11, and 13 are not functions of time. Thus, $x_t z_t$ will be stationary if and only if $\mathbb{C}[\hat{x}, \hat{z}] = \mathbb{C}[\tilde{x}, \tilde{z}]\ \forall\ \hat{x}$ and $\tilde{x} \in \{x_t, x_{t+\ell}, x_t x_{t+\ell}\}$, $\hat{z}$ and $\tilde{z} \in \{z_t, z_{t+\ell}, z_t z_{t+\ell}\}$, where $t, \ell \in \mathbb{Z}_{\geq 0}$, and $\hat{x}$ ($\hat{z}$) and $\tilde{x}$ ($\tilde{z}$) differ only in t. ∎

1.2 Corollary 1

Proof. Assume x and z are stochastically independent. I calculate the the mean, variance, and auto-covariance of $x_t z_t$ in turn.

Mean. The mean is given by

$$\begin{aligned}\mathbb{E}[x_t z_t] &= \mathbb{E}[x_t]\mathbb{E}[z_t] \\ &= \mu_x\mu_z.\end{aligned} \tag{14}$$

Variance. The variance is given by

$$\begin{aligned}\mathbb{V}[x_t z_t] &= (\mathbb{E}[x_t])^2\,\mathbb{V}(z_t) + (\mathbb{E}[z_t])^2\,\mathbb{V}(x_t) + \mathbb{V}(z_t)\,\mathbb{V}(x_t) \\ &= \mu_x^2\sigma_z^2 + \mu_z^2\sigma_x^2 + \sigma_x^2\sigma_z^2.\end{aligned} \tag{15}$$

Autocovariance. For independent x and z, all of the third moments in Equation 12 reduce to zero, and the fourth moment reduces to the product of each variable's autocovariance, such that

$$\begin{aligned}\mathbb{C}[x_t z_t, x_{t+\ell} z_{t+\ell}] &= \mathbb{E}\left[x_t\right]\mathbb{E}\left[x_{t+\ell}\right]\mathbb{C}\left[z_t, z_{t+\ell}\right] + \mathbb{E}\left[x_t\right]\mathbb{E}\left[z_{t+\ell}\right]\mathbb{C}\left[x_{t+\ell}, z_t\right] \\ &\quad + \mathbb{E}\left[z_t\right]\mathbb{E}\left[x_{t+\ell}\right]\mathbb{C}\left[x_t, z_{t+\ell}\right] + \mathbb{E}\left[z_t\right]\mathbb{E}\left[z_{t+\ell}\right]\mathbb{C}\left[x_t, x_{t+\ell}\right] \\ &\quad + \rho_x^\ell \rho_z^\ell \sigma_x^2 \sigma_z^2 - \mathbb{C}\left[x_t, z_t\right]\mathbb{C}\left[x_{t+\ell}, z_{t+\ell}\right] \\ &= \mu_x^2 \rho_z^\ell \sigma_z^2 + \mu_z^2 \rho_x^\ell \sigma_x^2 + \rho_x^\ell \rho_z^\ell \sigma_x^2 \sigma_z^2. \end{aligned} \tag{16}$$

Checking for stationarity. The unconditional mean is constant, all of the terms in the unconditional variance are positive constants, and all of the terms in the autocovariance are constant. Thus the stochastic process xz is stationary and can be described:

$$\begin{aligned} xz &\sim \mathcal{F}\left(\mu_x \mu_z, \mu_x^2 \sigma_z^2 + \mu_z^2 \sigma_x^2 + \sigma_x^2 \sigma_z^2\right), \\ \gamma_{xz}\left(\ell\right) &= \mu_x^2 \rho_z^\ell \sigma_z^2 + \mu_z^2 \rho_x^\ell \sigma_x^2 + \rho_x^\ell \rho_z^\ell \sigma_x^2 \sigma_z^2. \end{aligned} \qquad \blacksquare$$

1.3 Proposition 2

Proof. Recall that a cointegrating vector is a vector of coefficients such that the linear combination of more than one stochastic series produces a stationary stochastic process, where at least one element of the vector is non-zero. Assume y, x, and z are (potentially non-stationary) AR(1) stochastic processes. Define the trivariate system with no multiplicative interactions among covariates x and z as

$$y_t - \alpha_0 - \alpha_1 y_{t-1} - \beta_0 x_t - \beta_1 x_{t-1} - \beta_2 z_t - \beta_3 z_{t-1} = \epsilon_t, \tag{17}$$

where $\epsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$. Assume that there is at least one cointegrating vector for this system; without loss of generality, define this cointegrating vector in the normalized form $\boldsymbol{\beta} = (1, -\alpha_1, -\beta_0, -\beta_1, -\beta_2, -\beta_3)$.

Define the system that includes a multiplicative interaction between x_t and z_t as

$$y_t - \alpha_0 - \alpha_1 y_{t-1} - \beta_0 x_t - \beta_1 x_{t-1} - \beta_2 z_t - \beta_3 z_{t-1} - \beta_4 x_t z_t = \epsilon_t. \tag{18}$$

Consider the candidate cointegrating vector

$$\boldsymbol{\beta}' = (1, -\alpha_1, -\beta_0, -\beta_1, -\beta_2, -\beta_3, -\beta_4),$$

where $\beta_4 = 0$. It is evident that this produces a stationary stochastic process ϵ, by the assumption that $\boldsymbol{\beta}$ does: β_4 makes the multiplicative interaction zero, and by assumption, the remaining coefficients produce a linear combination of y, x, and z such that ϵ is stationary. Further, the assumption that $\boldsymbol{\beta}$ is a cointegrating vector implies that at least one term in $\boldsymbol{\beta}$—and therefore $\boldsymbol{\beta}'$—is non-zero. Thus $\boldsymbol{\beta}'$ is a cointegrating vector, and the system with a multiplicative interaction among covariates is cointegrated. ■

2 Extension to an AR(p) process

Proof. The response y and covariates x and z can be defined as stationary AR(p) processes:

$$y_t = \zeta_y + \rho_{y,1} y_{t-1} + \rho_{y,2} y_{t-2} + \cdots + \rho_{y,p} y_{t-p} + \eta_t \tag{19}$$
$$x_t = \zeta_x + \rho_{x,1} x_{t-1} + \rho_{x,2} x_{t-2} + \cdots + \rho_{x,p} x_{t-p} + \eta_t \tag{20}$$
$$z_t = \zeta_z + \rho_{z,1} z_{t-1} + \rho_{z,2} z_{t-2} + \cdots + \rho_{z,p} z_{t-p} + \eta_t. \tag{21}$$

Allow L to denote the lag operator—when L multiplies any parameter with a time subscript, this subscript is lagged one period—and define

$$\phi(L) = 1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p. \tag{22}$$

Stationarity requires that the roots of the polynomial equation $\phi(L) = 0$ lie outside the unit circle for y, x, and z. This allows us to rewrite the unconditional mean, unconditional variance, and autocovariance conditional on ℓ as

$$\mathbb{E}[y_t] = \frac{\zeta_y}{1 - \sum_{i=1}^{p} \rho_{y,i}}, \quad \mathbb{V}[y_t] = \frac{\sigma^2}{1 - \sum_{i=1}^{p} \rho_{y,i}^2}, \quad \mathbb{C}[y_t, y_{t+\ell}] = \sum_{r=1}^{p} \rho_{y,r} \gamma_y(\ell - r) \tag{23}$$

$$\mathbb{E}[x_t] = \frac{\zeta_x}{1 - \sum_{i=1}^{p} \rho_{x,i}}, \quad \mathbb{V}[x_t] = \frac{\sigma^2}{1 - \sum_{i=1}^{p} \rho_{x,i}^2}, \quad \mathbb{C}[x_t, x_{t+\ell}] = \sum_{r=1}^{p} \rho_{x,r} \gamma_x(\ell - r) \tag{24}$$

$$\mathbb{E}[z_t] = \frac{\zeta_z}{1 - \sum_{i=1}^{p} \rho_{z,i}}, \quad \mathbb{V}[z_t] = \frac{\sigma^2}{1 - \sum_{i=1}^{p} \rho_{z,i}^2}, \quad \mathbb{C}[z_t, z_{t+\ell}] = \sum_{r=1}^{p} \rho_{z,r} \gamma_z(\ell - r) \tag{25}$$

As above, define the mean, variance, and autocorrelation of y as μ_y, σ_y^2, and $\gamma_y(\ell)$, respectively, and similarly for x and z. The same result as in the AR(1) case holds. ■

3 Quantities of interest in the ECM framework

Recall that the general model with a conditional relationship is given in ADL form in Equations 2 and 3 (in the main text) in ADL and ECM form, respectively. All of the quantities described in the text can be recovered from either specification; Table 1 provides the variable translations to move between models. These can be used to rewrite the general equations for quantities of interest from the ECM, which I provide here.

Period-specific effects: Equation 4 (main text) becomes

$$\frac{\partial y_{t+j}}{\partial x_t} = \begin{cases} 0 & \text{for } j \in \mathbb{Z}_{<0}, \\ \theta_0 + \theta_4 z_t + (\theta_5 + \theta_6)\,\Delta z_t & \text{for } j = 0, \\ (\gamma_1 + 1)^j \left[\theta_0 + \theta_4 z_t + (\theta_5 + \theta_6)\,\Delta z_t\right] + \\ \qquad (\gamma_1 + 1)^{j-1} \left[\theta_1 - \theta_0 - \theta_4 z_{t+1} - \theta_6 \Delta z_{t+1} + \theta_7 z_t\right] & \text{for } j \in \mathbb{Z}_{>0}. \end{cases} \tag{26}$$

Cumulative effects: Equation 5 (main text) yields

$$\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = \frac{(\theta_0 + \theta_4 z_t + \theta_5 \Delta z_t + \theta_6 \Delta z_t)\left([\gamma_1 + 1]^h - [\gamma_1 + 1]^{k+1}\right)}{-\gamma_1} + \frac{(\theta_1 - \theta_0 - \theta_4 z_{t+1} - \theta_6 \Delta z_{t+1} + \theta_7 z_t)\left([\gamma_1 + 1]^{h-1} - [\gamma_1 + 1]^k\right)}{-\gamma_1}$$

for $h, k \in \mathbb{Z}_{>0}$ such that $h < k$, and

$$\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = \frac{(\theta_0 + \theta_4 z_t + \theta_5 \Delta z_t + \theta_6 \Delta z_t)\left(1 - [\gamma_1 + 1]^{k+1}\right)}{-\gamma_1} + \frac{(\theta_1 - \theta_0 - \theta_4 z_{t+1} - \theta_6 \Delta z_{t+1} + \theta_7 z_t)\left(1 - [\gamma_1 + 1]^k\right)}{-\gamma_1}$$

Table 1: Variable transformations between the general ADL and ECM in Equations 2 and 3 (main text)

ADL	ECM
$\alpha_0 = \gamma 0$	$\alpha_0 = \gamma_0$
$\alpha_1 = \gamma_1 + 1$	$\alpha_1 - 1 = \gamma_1$
$\beta_0 = \theta_0$	$\beta_0 = \theta_0$
$\beta_1 = \theta_1 - \theta_0$	$\beta_0 + \beta_1 = \theta_1$
$\beta_2 = \theta_2$	$\beta_2 = \theta_2$
$\beta_3 = \theta_3 - \theta_2$	$\beta_2 + \beta_3 = \theta_3$
$\beta_4 = \theta_4 + \theta_5 + \theta_6$	$\beta_4 - \beta_6 = \theta_4$
$\beta_5 = -(\theta_4 + \theta_6)$	$\beta_4 - \beta_5 = \theta_5$
$\beta_6 = -(\theta_5 + \theta_6)$	$-(\beta_4 + \beta_5 + \beta_6) = \theta_6$
$\beta_7 = \theta_6 + \theta_7$	$\beta_4 + \beta_5 + \beta_6 + \beta_7 = \theta_7$

for $h = 0$. Also note that the LRE is given by

$$\sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial x_t} = \frac{\theta_1 + \theta_7 z_t + (\theta_5 + \theta_6)\,\Delta z_t - (\theta_4 + \theta_6)\,\Delta z_{t+1}}{-\gamma_1}. \tag{27}$$

Threshold effects: First define

$$\begin{aligned} \mathbf{a} &= \theta_0 + \theta_4 z_t + (\theta_5 + \theta_6)\,\Delta z_t \\ \mathbf{b} &= \theta_1 - \theta_0 - \theta_4 z_{t+1} - \theta_6 \Delta z_{t+1} + \theta_7 z_t. \end{aligned}$$

for ease of presentation. Also define an arbitrary threshold λ, and solve for k from the cumulative effect where $h = 0$. This yields

$$k = \left\lceil \frac{\log\left(\dfrac{\mathbf{a} + \mathbf{b} - (1 - \alpha_1)\,\lambda}{\alpha_1 \mathbf{a} + \mathbf{b}}\right)}{\log \alpha_1} \right\rceil,$$

where the ceiling function indicates rounding up since k is a discrete period in time. Lag lengths are identically derived, where $\lambda \equiv \delta\Omega$, $\delta \in [0, 1)$, and Ω is the LRE. As discussed in the text, both quantities are more easily found inductively than solved for.

4 Quantities of interest

4.1 Period-specific effects

Recall that the general conditional model is given by

$$
\begin{aligned}
y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 z_t + \beta_3 z_{t-1} \\
+ \beta_4 x_t z_t + \beta_5 x_{t-1} z_t + \beta_6 x_t z_{t-1} + \beta_7 x_{t-1} z_{t-1} + \epsilon_t.
\end{aligned} \tag{28}
$$

Period-specific effects are defined as $\frac{\partial y_{t+j}}{\partial x_t}\ \forall\ j \in \mathbb{Z}$. We first examine the case where $j = 0$, which is just Equation 28. Differentiating with respect to x_t yields

$$
\frac{\partial y_t}{\partial x_t} = \beta_0 + \beta_4 z_t + \beta_6 z_{t-1}.
$$

Moving to $j = 1$, the model becomes

$$
\begin{aligned}
y_{t+1} = \alpha_0 + \alpha_1 y_t + \beta_0 x_{t+1} + \beta_1 x_t + \beta_2 z_{t+1} + \beta_3 z_t \\
+ \beta_4 x_{t+1} z_{t+1} + \beta_5 x_t z_{t+1} + \beta_6 x_{t+1} z_t + \beta_7 x_t z_t + \epsilon_{t+1},
\end{aligned}
$$

into which we can substitute the right side of Equation 28 for y_t. This yields

$$
\begin{aligned}
y_{t+1} = \alpha_0 + \alpha_1 (\alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 z_t + \beta_3 \\
+ z_{t-1} + \beta_4 x_t z_t + \beta_5 x_{t-1} z_t + \beta_6 x_t z_{t-1} + \beta_7 x_{t-1} z_{t-1} + \epsilon_t) \\
+ \beta_0 x_{t+1} + \beta_1 x_t + \beta_2 z_{t+1} + \beta_3 z_t \\
+ \beta_4 x_{t+1} z_{t+1} + \beta_5 x_t z_{t+1} + \beta_6 x_{t+1} z_t + \beta_7 x_t z_t + \epsilon_{t+1}.
\end{aligned}
$$

Differentiating yields

$$
\frac{\partial y_{t+1}}{\partial x_t} = \alpha_1 (\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}) + \beta_1 + \beta_5 z_{t+1} + \beta_7 z_t,
$$

which is the period-specific effect for $j = 1$. Following the same process for $j = 2$ gives $\frac{\partial y_{t+2}}{\partial x_t} = \alpha_1^2 (\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}) + \alpha_1 (\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t)$. This pattern continues for all $j \in \mathbb{Z}_{>0}$, with each set of terms in parantheses multiplied by α_1 to increasingly higher powers. Finally, note that for any $j < 0$, no substitutions can be made such that an x_t term appears among the covariates, making each derivative zero. Thus, for negative j, all period-specific effects are nil. This gives the general expression for period-specific effects (Equation 4 in the main text).

4.2 Cumulative effects

Cumulative effects are sums of period-specific effects. Defined most generally as

$$\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} \equiv \frac{\partial y_{t+h}}{\partial x_t} + \frac{\partial y_{t+h+1}}{\partial x_t} + \cdots + \frac{\partial y_{t+k}}{\partial x_t},$$

we can substitute such that

$$\begin{aligned}
\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = {} & \alpha_1^{h} \left(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}\right) + \alpha_1^{h-1} \left(\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t\right) \\
& + \alpha_1^{h+1} \left(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}\right) + \alpha_1^{h} \left(\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t\right) \\
& \vdots \\
& + \alpha_1^{k} \left(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}\right) + \alpha_1^{k-1} \left(\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t\right).
\end{aligned}$$

Group terms and rewrite:

$$\begin{aligned}
\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = {} & \left(\alpha_1^{h} + \alpha_1^{h+1} + \cdots + \alpha_1^{k}\right) \left(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}\right) \\
& + \left(\alpha_1^{h-1} + \alpha_1^{h} + \cdots + \alpha_1^{k-1}\right) \left(\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t\right). \qquad (29)
\end{aligned}$$

These are two geometric sums which can be solved simultaneously. Multiply both through by α_1 and then subtract the result from Equation 29, giving

$$(1-\alpha_1)\sum_{j=h}^{k}\frac{\partial y_{t+j}}{\partial x_t} = \left(\alpha_1^h - \alpha_1^{k+1}\right)(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}) \\ + \left(\alpha_1^{h-1} - \alpha_1^{k}\right)(\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t)\,.$$

Dividing through by $(1-\alpha_1)$ gives the expression for $0 < h < k$ (Equation 5 in the main text). Note, however, that when $h = 0$, this expression raises $\beta_1 + \beta_5 z_{t+1} + \beta_7 z_t$ to a negative power. This is not correct, and the instantaneous effect does not include any of these terms. Thus $h = 0$ presents a special case, for which repeating the same approach to derivation with $h = 0$ yields the general expression in Equation 6 in the main text.

4.3 Threshold effects

For ease of presentation, define

$$\mathbf{a} = \beta_0 + \beta_4 z_t + \beta_6 z_{t-1}$$
$$\mathbf{b} = \beta_1 + \beta_5 z_{t+1} + \beta_7 z_t.$$

Also define an arbitrary threshold λ, and solve for k in Equation 6 (main text) where $h = 0$. This yields

$$k = \left\lceil \frac{\log\left(\dfrac{\mathbf{a}+\mathbf{b}-(1-\alpha_1)\lambda}{\alpha_1\mathbf{a}+\mathbf{b}}\right)}{\log\alpha_1} \right\rceil,$$

where the ceiling function indicates rounding up since k is a discrete period in time. Lag lengths are identically derived, where $\lambda \equiv \delta\Omega$, $\delta \in [0,1)$, and Ω is the LRE. As discussed in the text, both quantities are more easily found inductively than solved for.

4.4 Estimating uncertainty with parametric bootstrapping

Building on King, Tomz, and Wittenberg (2000), we can describe the parametric bootstrap for a maximum likelihood framework:

1. Estimate the model, e.g. Equation 28, from which we wish to draw inferences.
2. Store the point estimates ($\hat{\boldsymbol{\beta}}$) and variance-covariance matrix ($\mathbb{C}(\hat{\boldsymbol{\beta}})$).
3. Take M draws from the multivariate normal $\mathcal{N}_{MV}\left(\hat{\boldsymbol{\beta}}, \mathbb{C}(\hat{\boldsymbol{\beta}})\right)$ and save the output.
4. Define S scenarios of interest to study. For instance, we may hold all variables at their central tendency but vary one along its interquartile range by S increments. With dynamic models, these scenarios need to satisfy the constraint $\Delta y_t = y_t - y_{t-1}$ (and similarly for x and z) to be empirically relevant.
5. For each $s \in S$ and $m \in M$, calculate quantities of interest.
6. For each $s \in S$, take quantiles over the M draws, e.g. the .025$^{\text{th}}$ and .975$^{\text{th}}$, to approximate 95% confidence intervals.

This procedure yields statements of uncertainty for every quantity of interest for every scenario $s \in S$. Example code for this procedure is provided in the replication archive.

5 Quantities of interest for a model with no interaction

The inferential framework developed for the paper is suitable for deriving quantities of interest from any ADL or ECM. To illustrate, I provide quantities of interest for the basic ADL(1,1;1) studied in De Boef and Keele (2008), given by:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t. \tag{30}$$

5.1 Period-specific effects

These are defined as $\frac{\partial y_{t+j}}{\partial x_t} \forall j \in \mathbb{Z}$. We first examine the case where $j = 0$. Differentiating with respect to x_t yields

$$\frac{\partial y_t}{\partial x_t} = \beta_0,$$

which scholars commonly refer to as the "short term effect" but I discuss as the *instantaneous effect* in the main text.

Moving to $j = 1$, the model becomes

$$y_{t+1} = \alpha_0 + \alpha_1 y_t + \beta_0 x_{t+1} + \beta_1 x_t + \epsilon_{t+1},$$

into which we can substitute the right side of Equation 30 for y_t. This yields

$$y_{t+1} = \alpha_0 + \alpha_1 \left(\alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t\right) + \beta_0 x_{t+1} + \beta_1 x_t + \epsilon_{t+1}.$$

Differentiating yields

$$\frac{\partial y_{t+1}}{\partial x_t} = \alpha_1 \left(\beta_0 + \beta_4 z_t + \beta_6 z_{t-1}\right) + \beta_1 + \beta_5 z_{t+1} + \beta_7 z_t,$$

which is the period-specific effect for $j = 1$. Following the same process for $j = 2$ gives $\frac{\partial y_{t+2}}{\partial x_t} = \alpha_1^2 \left(\beta_0\right) + \alpha_1 \left(\beta_1\right)$. This pattern continues for all $j \in \mathbb{Z}_{>0}$, with each set of terms in parantheses multiplied by α_1 to increasingly higher powers. Finally, note that for any $j < 0$, no substitutions can be made such that an x_t term appears among the covariates, making each derivative zero. Thus, for negative j, all period-specific effects are nil. This gives the following general expression:

$$\tfrac{\partial y_{t+j}}{\partial x_t} = \begin{cases} 0 & \text{for } j \in \mathbb{Z}_{<0}, \\ \beta_0 & \text{for } j = 0, \\ \alpha_1^j \left(\beta_0\right) + \alpha_1^{j-1} \left(\beta_1\right) & \text{for } j \in \mathbb{Z}_{>0}. \end{cases} \tag{31}$$

5.2 Cumulative effects

Cumulative effects are sums of period-specific effects. Defined most generally as

$$\sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} \equiv \frac{\partial y_{t+h}}{\partial x_t} + \frac{\partial y_{t+h+1}}{\partial x_t} + \cdots + \frac{\partial y_{t+k}}{\partial x_t},$$

we can substitute such that

$$\begin{aligned} \sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = {} & \alpha_1^h \left(\beta_0\right) + \alpha_1^{h-1} \left(\beta_1\right) \\ & + \alpha_1^{h+1} \left(\beta_0\right) + \alpha_1^{h} \left(\beta_1\right) \\ & \vdots \\ & + \alpha_1^{k} \left(\beta_0\right) + \alpha_1^{k-1} \left(\beta_1\right). \end{aligned}$$

Group terms and rewrite:

$$\begin{aligned} \sum_{j=h}^{k} \frac{\partial y_{t+j}}{\partial x_t} = {} & \left(\alpha_1^h + \alpha_1^{h+1} + \cdots + \alpha_1^{k}\right) \left(\beta_0\right) \\ & + \left(\alpha_1^{h-1} + \alpha_1^{h} + \cdots + \alpha_1^{k-1}\right) \left(\beta_1\right). \end{aligned} \tag{32}$$

These are two geometric sums which can be solved simultaneously. Multiply both through by α_1 and then subtract the result from Equation 32, giving

$$(1-\alpha_1)\sum_{j=h}^{k}\frac{\partial y_{t+j}}{\partial x_t} = \left(\alpha_1^h - \alpha_1^{k+1}\right)(\beta_0) + \left(\alpha_1^{h-1} - \alpha_1^k\right)(\beta_1).$$

Dividing through by $(1-\alpha_1)$ gives the following general expression for $0 < h < k$:

$$\sum_{j=h}^{k}\tfrac{\partial y_{t+j}}{\partial x_t} = \frac{(\beta_0)\left(\alpha_1^h - \alpha_1^{k+1}\right) + (\beta_1)\left(\alpha_1^{h-1} - \alpha_1^k\right)}{1-\alpha_1}. \tag{33}$$

Again $h = 0$ presents a special case, given by

$$\sum_{j=0}^{k}\tfrac{\partial y_{t+j}}{\partial x_t} = \frac{(\beta_0)\left(1 - \alpha_1^{k+1}\right) + (\beta_1)\left(1 - \alpha_1^k\right)}{1-\alpha_1}. \tag{34}$$

Finally, if we allow $k \to \infty$, we get the total effect:

$$\sum_{j=0}^{\infty}\tfrac{\partial y_{t+j}}{\partial x_t} = \frac{\beta_0 + \beta_1}{1-\alpha_1}. \tag{35}$$

This quantity is clearly the widely-known LRE (or "LRM") for an ADL(1,1;1).

5.3 Threshold effects

Define an arbitrary threshold λ, and solve for k in Equation 34. This yields

$$k = \left\lceil \frac{\log\left(\frac{\beta_0 + \beta_1 - (1-\alpha_1)\lambda}{\alpha_1\beta_0 + \beta_1}\right)}{\log \alpha_1} \right\rceil,$$

where the ceiling function indicates rounding up since k is a discrete period in time. Lag lengths are identically derived, where $\lambda \equiv \delta\Omega$, $\delta \in [0,1)$, and Ω is the LRE. As discussed in the text, both quantities are more easily found inductively than solved for.

6 Further Monte Carlo results

In this section I provide further Monte Carlo results. Each figure plots mean estimates in the top panel and coverage rates in the bottom panel, analogous to Figure 1 for the parameters of interest not discussed in the text. Like with the error-correction rate, I find that the general model produces better estimates of *all* parameters than does a model with an invalid restriction.

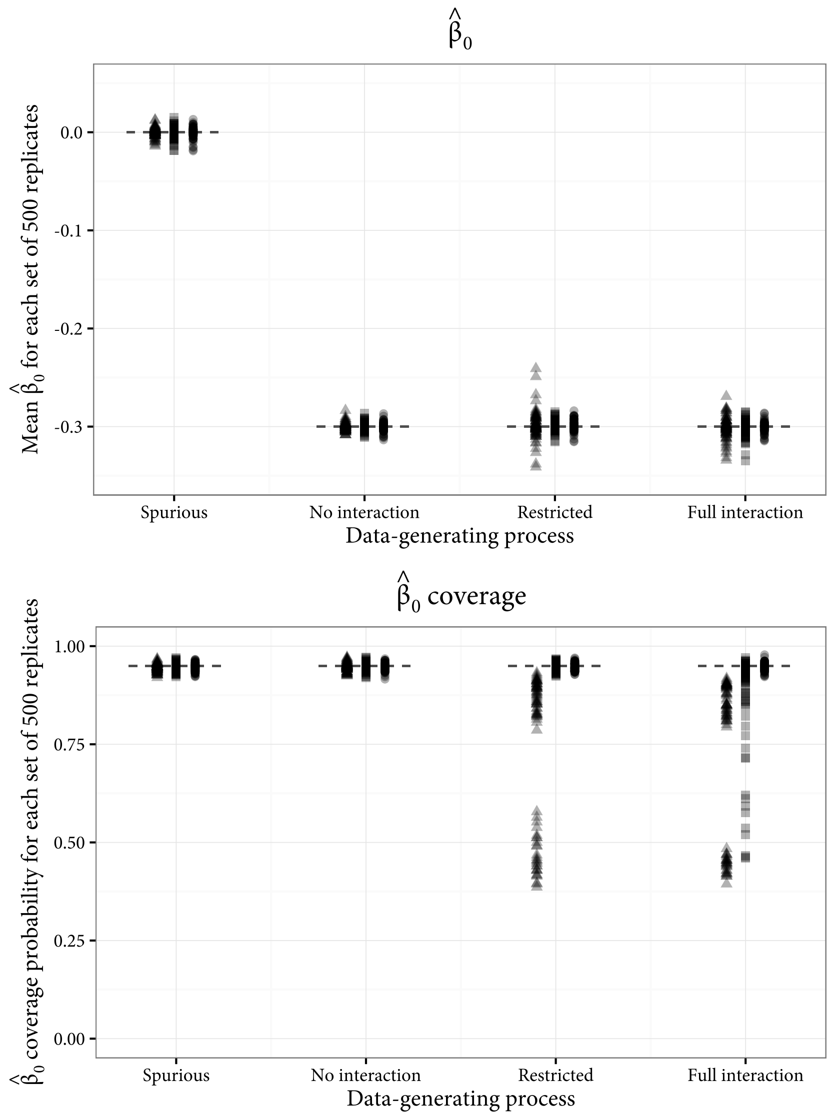

Figure 1: Monte Carlo simulation results for β_0. Triangles represent estimates from the model without an interaction, squares are from a restricted model, and circles are from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

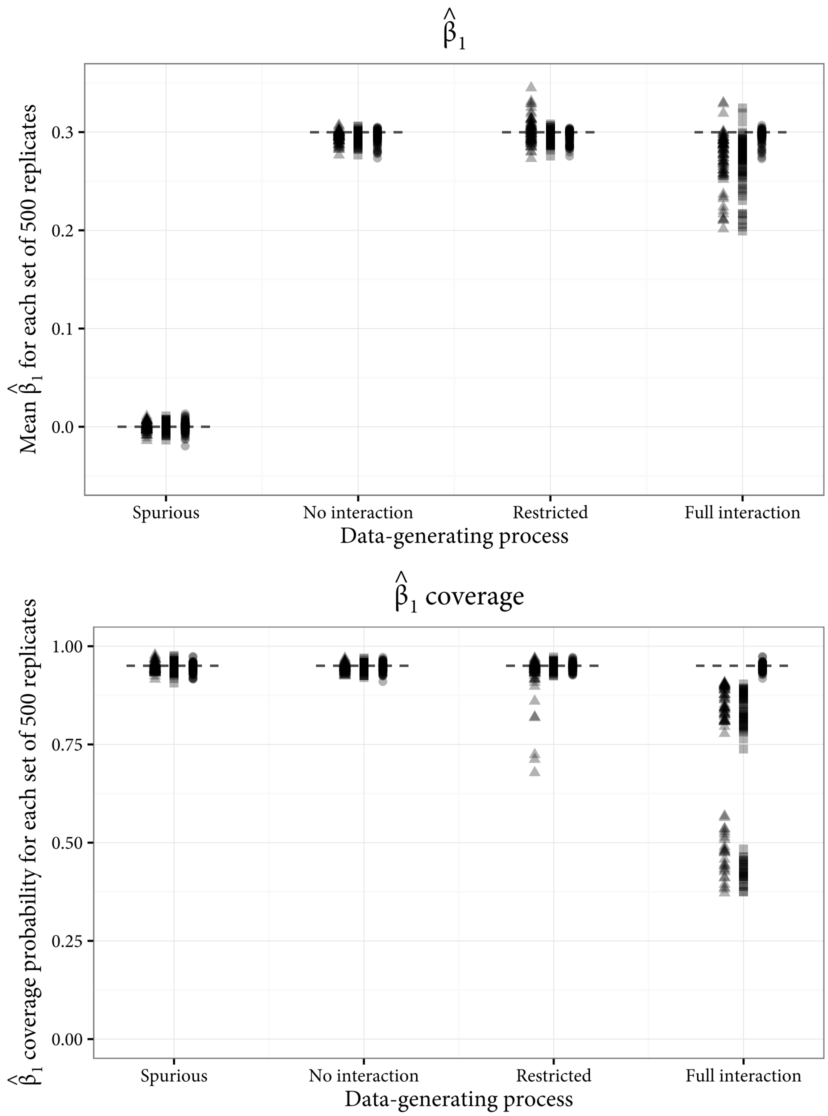

Figure 2: Monte Carlo simulation results for β_1. Triangles represent estimates from the model without an interaction, squares are from a restricted model, and circles are from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

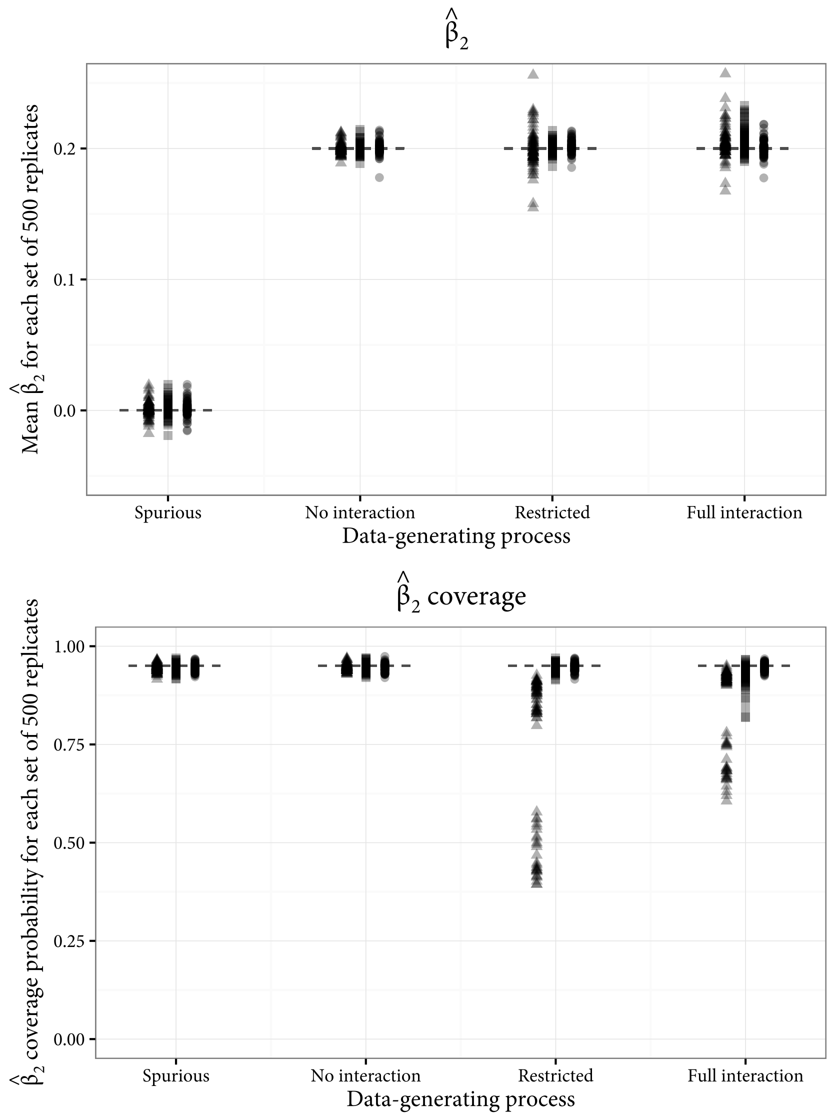

Figure 3: Monte Carlo simulation results for β_2. Triangles represent estimates from the model without an interaction, squares are from a restricted model, and circles are from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

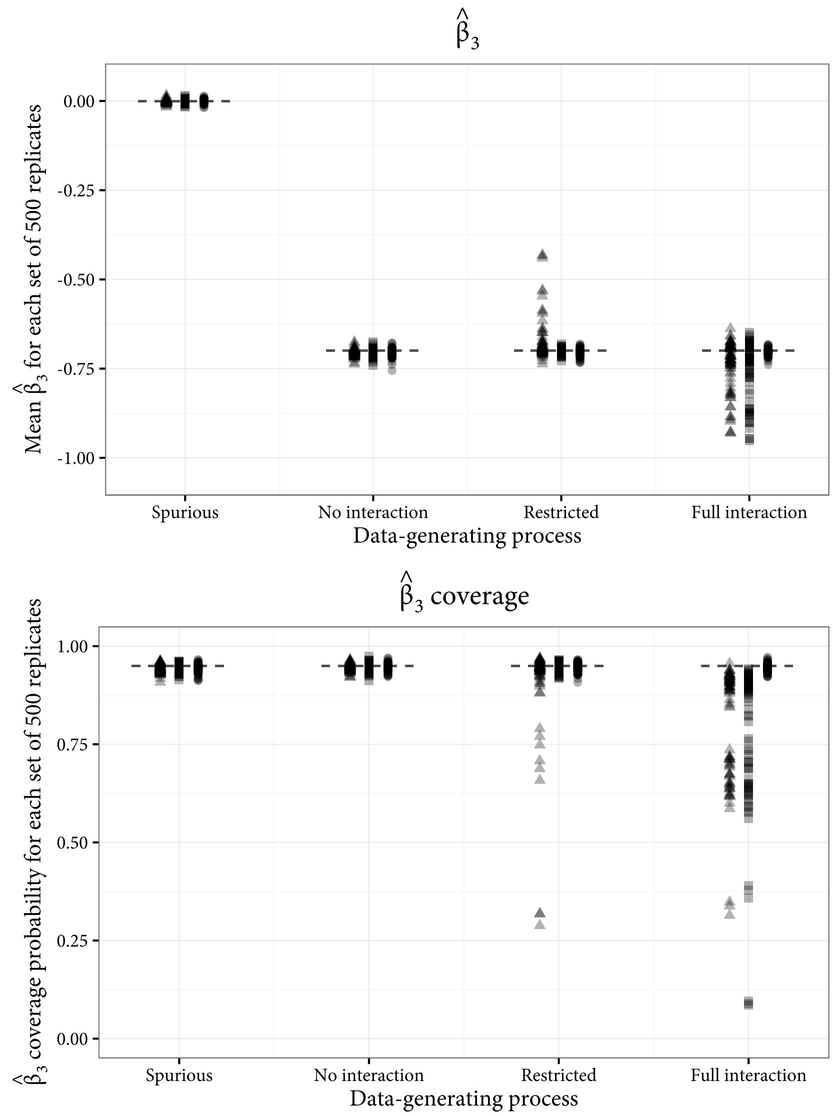

Figure 4: Monte Carlo simulation results for β_3. Triangles represent estimates from the model without an interaction, squares are from a restricted model, and circles are from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

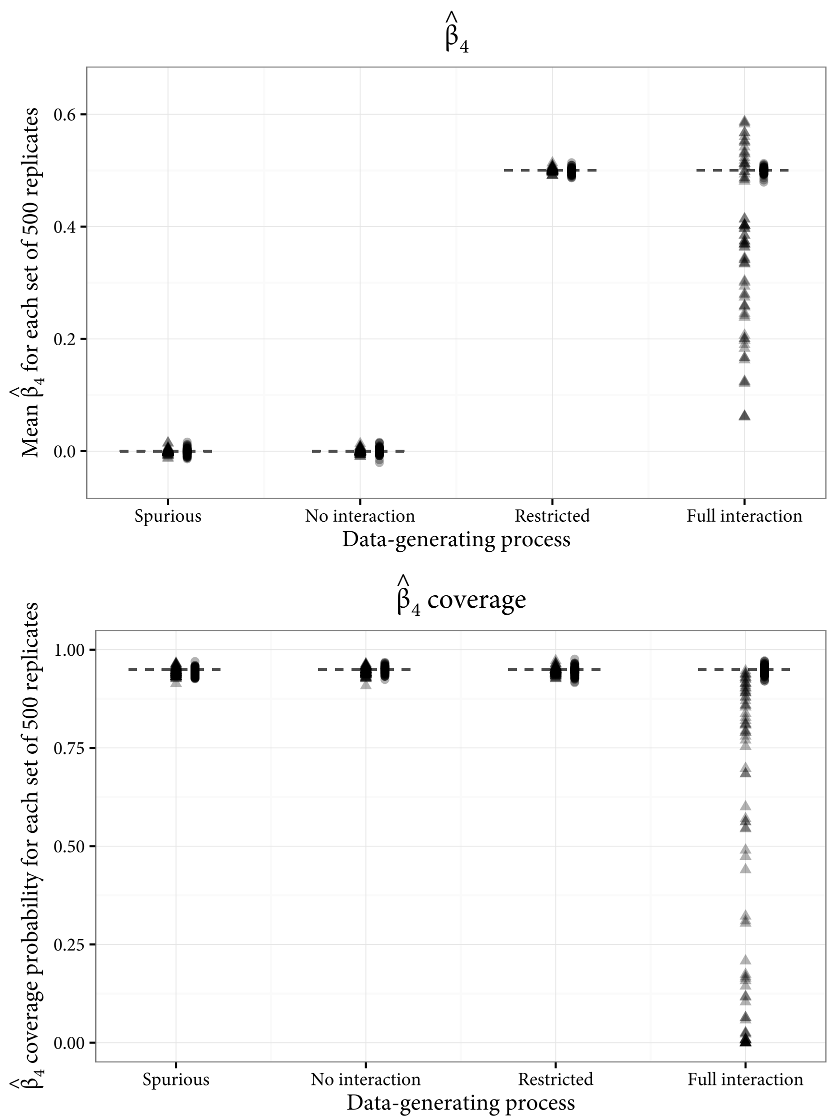

Figure 5: Monte Carlo simulation results for β_4. Squares represent estimates from a restricted model and circles are from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

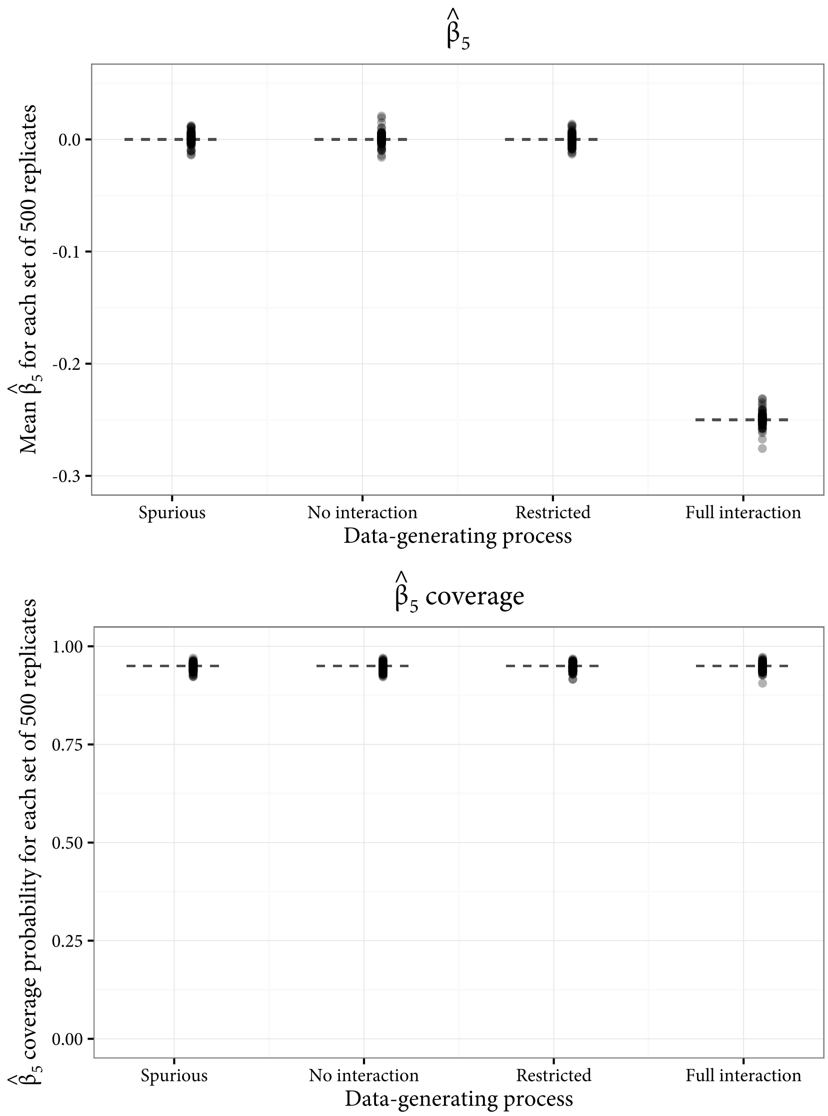

Figure 6: Monte Carlo simulation results for β_5. Circles represent estimates from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

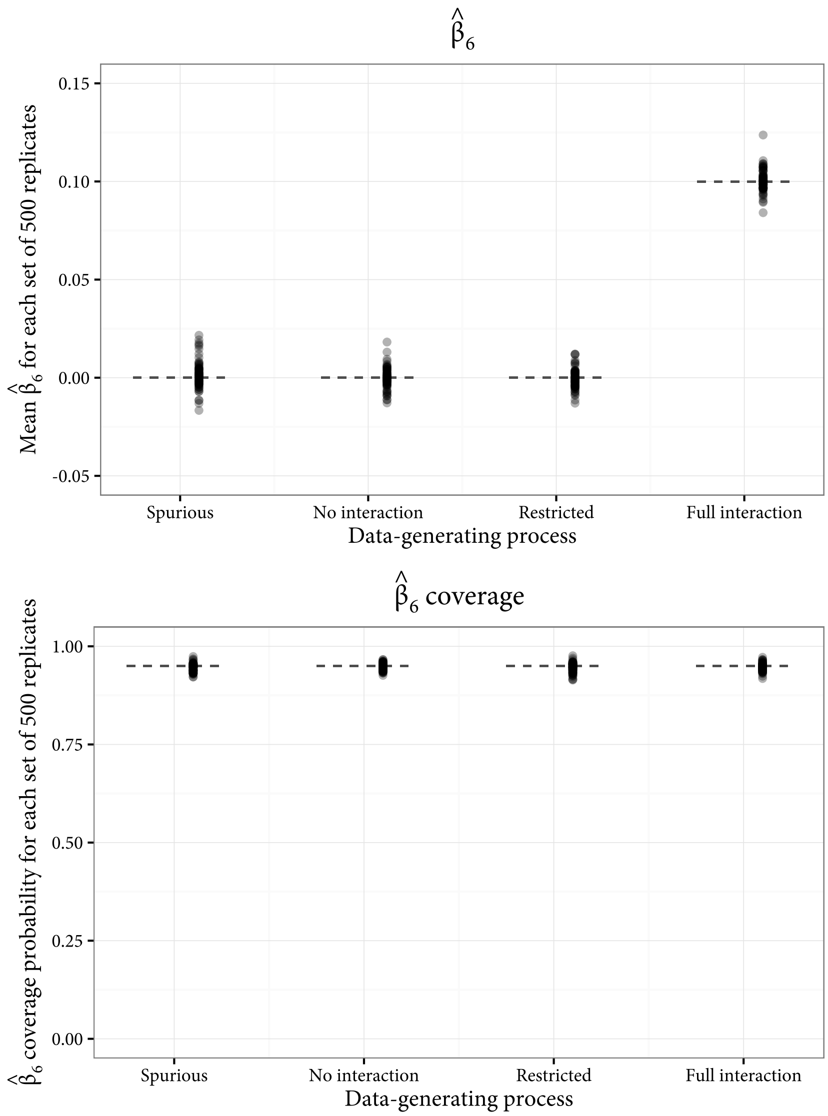

Figure 7: Monte Carlo simulation results for β_6. Circles represent estimates from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

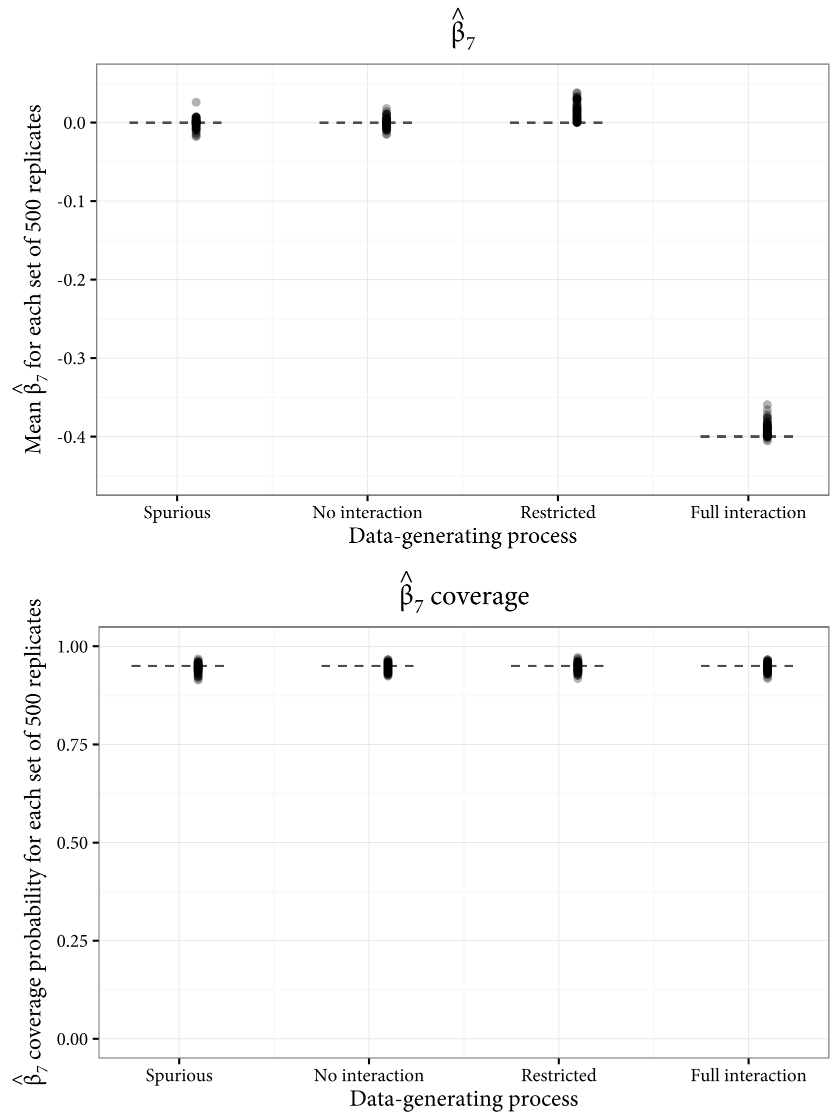

Figure 8: Monte Carlo simulation results for β_7. Circles represent estimates from the general model. Mean estimates are plotted against the true values (dashed lines) in the top panel. The bottom panel plots the proportion of 500 replicates under each condition for which 95% confidence intervals contain the true value.

7 Replication of Morgan and Kelly (2013)

How can states ensure that the poor benefit from growth? Many studies have uncovered mechanisms through which governments directly redistribute wealth over both the short and long run (e.g., Morley 2001), including such channels as explicit cash transfers and investments in primary education. Yet very little research is devoted to how states indirectly induce distributive outcomes by shaping market forces.

Morgan and Kelly (2013, "M & K") investigate these "market conditioning" mechanisms. They note how certain laws and regulations change the economic incentives of private actors—for instance, by inducing monetary stability or encouraging employment of low-wage workers—with knock-on effects for inequality. M & K argue that these market conditioning effects are greatest for human capital spending (HCS), which enhances the skills and productivity of the poor, increasing the value of their labor. Thus, M & K hypothesize that when HCS is low, the rich reap the gains from growth and inequality rises, but when it is high, growth is more efficiently distributed to the poor and inequality declines.

7.1 Data and estimation

Data for the response come from the Standardized World Income Inequality Database (Solt 2009). M & K's main analysis is based on gross adult-equivalent income adjusted for household composition, which captures pre-tax and -transfer income. From this the authors construct a market Gini index that ranges over 0-100, where larger values indicate a less equal distribution of wealth. For the covariates, M & K use economic growth in real per capita gross domestic product (GDP), measured in thousands of US dollars adjusted for purchasing power parity, and a fifteen-year average of aggregate HCS as a percentage of GDP, both from the Penn World Tables (Heston, Summers, and Aten 2001). Due to data constraints, the sample is restricted to 19 Latin American and Caribbean countries over the period 1980-2000.

Maintaining notation consistent with the general model in Equation 3 in the main text, the authors' preferred specification is the ECM

$$\Delta y_t = \gamma_0 + \gamma_1 y_{t-1} + \theta_0 \Delta x_t + \theta_2 \Delta z_t + \theta_3 z_{t-1} + \theta_4 \Delta x_t z_{t-1} + \epsilon_t,$$

where y is market Gini, x is GDP, z is HCS, and ϵ is the error term. Note that this model restricts $\theta_1 = \theta_5 = \theta_6 = \theta_7 = 0$, or equivalently, it translates to an ADL where $\beta_1 = -\beta_0$, $\beta_4 = \beta_5 = 0$, and $\beta_7 = -\beta_6$. Using the approach developed above, I solve for the LRE of a shock to GDP, which yields $\frac{-\theta_4 \Delta z_{t+1}}{-\gamma_1}$ (or $\frac{\beta_7 \Delta z_t}{1-\alpha_1}$). Thus, the authors' specification translates into a dynamic system in which both GDP and human capital spending must be *simultaneously* shocked for there to be any conditional relationship over the infinitely long run. This description of the dynamic system contrasts with their theoretical discussion, which predicts that "the *level* of investment in human capital will condition the effect of economic *growth*, or change in GDP" (Morgan and Kelly 2013, 679, italics in original). Moreover, this specification implies that if such simultaneous shocks do not occur, then the long-run effect of growth on regime change will be zero. These unintuitive implications of the M & K specification follow directly from the parameter restrictions imposed.

I replicate the M & K model before estimating the general model (Equation 3 in the main text). To ensure comparability, I include the control variables from M & K's favored specification in both.[1] I also use the authors' data, which centers all variables before estimation, and follow them in clustering standard errors by country.

M & K conduct extensive pretesting of their data, using a mix of unit root and cointegration tests. I extend this pretesting using the same mix of formal tests and visual diagnostics discussed in the paper. Again pretesting is complicated by the panel structure, but the evidence indicates that the interaction is a unit root process—in line with both intuition and other analyses of GDP and spending (see, e.g., Box-Steffensmeier et al. 2014). Thus, it is both integrated and—by Proposition 2—cointegrated with the other variables, which I confirm using the Engle-Granger and Johansen cointegration tests. The evidence indicates that the ECM is an appropriate model for these data.

1. These include both lags and differences of legislative partisan balance, inflation, unemployment, and foreign direct investment.

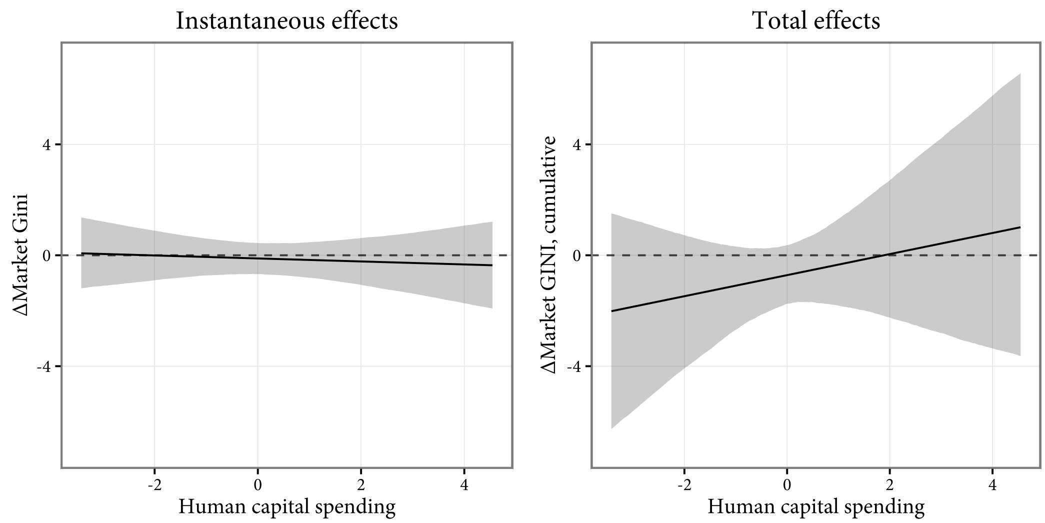

Figure 9: Short- and long-run relationships between growth and inequality, conditional on human capital spending. Lines represent predicted instantaneous effects (left) and total effects (right) on market Gini of mean growth in per capita GDP, holding growth in HCS at its mean value, while varying HCS across its observed range. 95% confidence intervals constructed from quantiles of 5,000 samples from $\mathcal{N}_{MV}\left(\hat{\boldsymbol{\theta}}, \mathbb{C}(\hat{\boldsymbol{\theta}})\right)$ are shaded in gray.

7.2 Results

Results for both models are presented in Table 2. Column 2 reproduces the authors' headline findings and column 3 reports estimates for the general model. Most notable among these estimates is that the main finding in M & K disappears in the full model: there appears to be no conditional relationship between growth and HCS on inequality. This result suggests that their model ignores the complex "memory" of the interaction term, inducing bias. Once other cross-time interactions are modeled, only HCS continues to exert a statistically significant effect.

Figure 9 visualizes this result. The first panel plots instantaneous effects of economic growth equal to the sample mean (approximately $\$63$ per capita after the authors' centering). Varying the level of HCS across its observed range while holding change in HCS at its mean, we see that growth never has an effect distinguishable from zero. The

Table 2: Replication of Morgan and Kelly (2013): comparing the restricted and general models

	M & K model	General model
Market Gini_{t-1}	-0.07^{*}	-0.08^{*}
	(0.02)	(0.02)
ΔGDP_t	-0.04	-0.19
	(0.15)	(0.26)
GDP_{t-1}		-0.03
		(0.04)
ΔHCS_t	0.74	2.68^{*}
	(1.28)	(1.90)
HCS_{t-1}	-0.16^{*}	-0.19^{*}
	(0.06)	(0.06)
$\Delta\text{GDP}_t \times \text{HCS}_{t-1}$	-0.24^{*}	-0.06
	(0.11)	(0.12)
$\text{GDP}_{t-1} \times \Delta\text{HCS}_t$		-0.73
		(0.56)
$\Delta\text{GDP}_t \times \Delta\text{HCS}_t$		2.39
		(2.14)
$\text{GDP}_{t-1} \times \text{HCS}_{t-1}$		0.03
		(0.04)
Observations	197	197
R^2	0.25	0.28
RMSE	1.01	1.01

$^{*}p < .05$. The dependent variable is ΔMarket Gini. Control variables are not reported and standard errors are clustered by country. RMSE is calculated from out-of-sample predictions using 5-fold cross-validation.

second panel plots LREs for the same scenarios; total effects are also indistinguishable from zero. Together, these results suggest that growth does not have an effect on inequality in either the short or long run, irrespective of the level of human capital spending.

Other quantities tell a similar story. Figure 10 plots cumulative effects over time, with both the level of and change in HCS fixed at their mean values. At no point in the 20 years following economic growth is there a discernible effect. Moreover, since

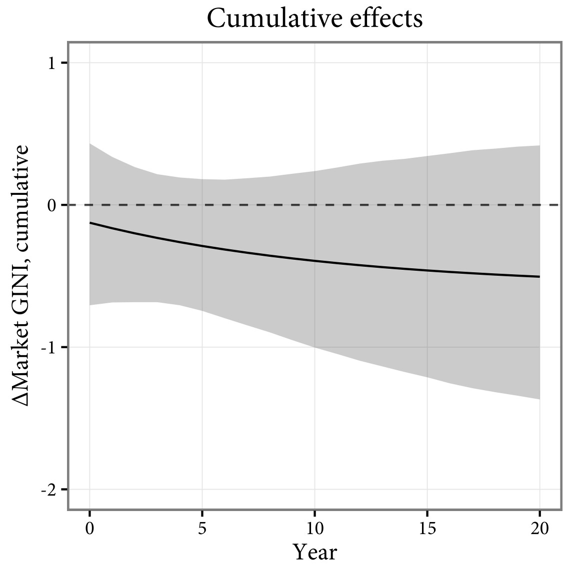

Figure 10: Dynamic relationship between growth and inequality, conditional on human capital spending. The line represents estimates of cumulative change in market Gini over time as the result of mean growth in per capita GDP, holding growth in (and level of) HCS at its mean value. 95% confidence intervals constructed from quantiles of 5,000 samples from $\mathcal{N}_{MV}\left(\hat{\boldsymbol{\theta}}, \mathbb{C}(\hat{\boldsymbol{\theta}})\right)$ are shaded in gray.

cumulative effects overlap across the entire window, none of the period-specific effects (not plotted) can be distinguished from zero. In short, we cannot be sure of any effect within or across any temporal windows.

Together these findings significantly revise the conclusions in Morgan and Kelly (2013): there is little evidence to support the market-conditioning hypothesis. Yet they also suggest a more subtle relationship between investment in human capital and inequality. While total effects for the mean change in HCS are negative and significant in poor countries, predicted instantaneous effects are positive. Thus, although greater social spending is associated with less inequality in the long run, it actually induces *greater* inequality in the short run. These results suggest that increased HCS comes at a cost which, at least at first, falls on the poorest. Scholars may need to revisit theories of the relationship between redistributive programs and inequality, paying greater attention to the incidence of increased social spending. Whether these patterns are echoed in other contexts remains to be seen, but either way, inequality in Latin America and the Caribbean appears to be a function of simple redistribution, and not growth as mediated by market conditioning institutions.

8 Replication of Jennings and John (2009)

What is the relationship between public opinion and government attention to various policy priorities? Jennings and John (2009, "J & J") study the effect of shifts in public opinion on mentions of policy topics in the Queen's Speech, a formal statement of the legislative agenda in the United Kingdom (UK). They provide extensive evidence that the variables exist in a dynamic equilibrium. J & J then ask (in their Table 4) if the effect of public attention to macroeconomic issues on Queen's Speech mentions of economic policy (which I will refer to as simply QS) is mediated by real indicators such as inflation and unemployment.

8.1 Data and estimation

Data for the dependent variable come from the UK Policy Agendas Project, which codes the number of mentions of major policy issue areas at the quasi-sentence level for every Speech from the Throne over the period 1911-2012. These data are expressed as an absolute count rather than a percentage of mentions, and so are unbounded (upward). For the purposes of their analysis of the conditional relationship with real economic indicators, J & J focus on the issue area relating to macroeconomics (which they refer to as QS1).

The covariate of interest is the proportion of citizens who believe an economic issue is the most important problem (MIP) facing the UK. The data come from two sources: the King and Wybrow (2001) collection of surveys covering 1937-2000 and the Gallup (2001) original index of surveys over 1959-2001. J & J merge and standardize these data into a single series. To examine economic conditions, J & J study the "misery index," a simple amalgamation of the inflation and unemployment rates (source not specified).

Maintaining notation consistent with the general model (Equation 3 in the main text), the authors' specification with an interaction between MIP and misery is the ECM

$$\Delta y_t = \gamma_0 + \gamma_1 y_{t-1} + \theta_6 \Delta x_t \Delta z_t + \theta_7 x_{t-1} z_{t-1} + \epsilon_t,$$

where y is QS, x is MIP, z is misery, and ϵ is error. Note that this model restricts

$\theta_0 = \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = 0$, or equivalently, it translates to an ADL where $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$ and $\beta_4 = -\beta_5 = -\beta_6$. Here the effect of these constraints is to restrict the instantaneous effect of a shock to MIP on QS to $\theta_6 \Delta z_t$, while the LRE reduces to $\frac{\theta_7 z_t + \theta_6 \Delta z_t - \theta_6 \Delta z_{t+1}}{-\gamma_1}$. Together, these quantities suggest that public attentiveness to economic issues affects the legislative agenda *exclusively* through real economic conditions: none of the terms in either quantity is independent of levels of (or change in) the misery index. This implication of the model jars with their theoretical argument, which posits both a direct effect of public opinion on Queen's Speech policy mentions, separate from the interactive effect.

As above, I first replicate the authors' model before estimating the general model. I include J & J's only control variable, an indicator for which party is in government. One observation drops out of both models due to missingness, leaving a time series of just 41 observations.

J & J present a number of tests of residuals for all of the estimated ECMs. I build on this by pretesting each variable for unit roots and visually inspecting ACF and PACF plots. The evidence indicates that each series is integrated. I then examine residuals from the full model, as well as conduct Johansen and Engle-Granger two-step tests to confirm cointegration. While these data are not in a panel, making pretesting slightly more straightforward, the time series is relatively short, diminishing the power of such tests. Again, there is some mixed evidence, but on balance, it points toward the ECM being an appropriate model for these data.

8.2 Results

Table 3 presents the results for the J & J model in column 2 and the general model in column 3. Notably, all significant coefficients disappear (except for the control variable for partisanship, which is omitted in the Table). There appears to be neither direct nor conditional relationships between public sentiment, economic conditions, and legislative attention. These null results are illustrated in Figure 11. At no level of economic misery does an average shift in public attention yield any effect on legislative attention to economic issues.

Table 3: Replication of Jennings and John (2009): comparing the restricted and general models

	J & J model	General model
QS_{t-1}	-1.00^{*}	-0.93^{*}
	(0.17)	(0.21)
ΔMIP_t		3.21
		(14.76)
MIP_{t-1}		1.51
		(7.78)
ΔMisery_t		-0.38
		(0.83)
Misery_{t-1}		0.09
		(0.38)
$\Delta\text{MIP}_t \times \text{Misery}_{t-1}$		0.12
		(0.96)
$\text{MIP}_{t-1} \times \Delta\text{Misery}_t$		1.04
		(1.31)
$\Delta\text{MIP}_t \times \Delta\text{Misery}_t$	0.44^{*}	1.14
	(0.15)	(1.66)
$\text{MIP}_{t-1} \times \text{Misery}_{t-1}$	0.38^{*}	0.23
	(0.10)	(0.60)
Observations	41	41
R^2	0.49	0.51
RMSE	2.67	3.44

$^{*}p < .05$. The dependent variable is ΔQS. Control variables are not reported. RMSE is calculated from out-of-sample predictions using 5-fold cross-validation.

Given the Monte Carlo results above, it seems likely that the estimates from J & J's model suffer from some bias arising from mis-specification. Yet the small sample size, combined with the number of parameters in the general model, suggest that overfitting is a real risk for these data. This interpretation is substantiated by the higher out-of-sample predictive RMSE for the general model. In short, both models likely suffer from bias of some sort. The most reasonable conclusion appears to be that there is simply not enough information in the data to reliably estimate this interaction.

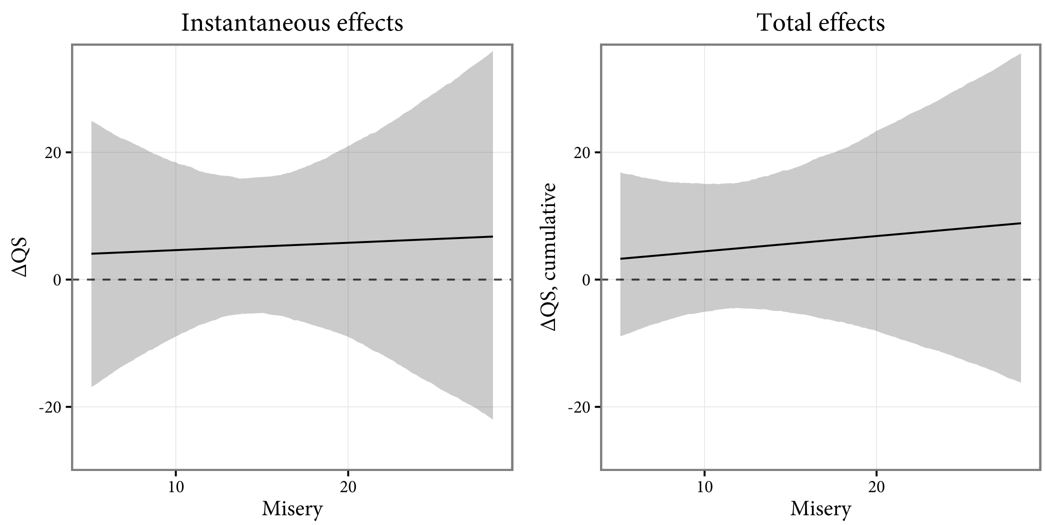

Figure 11: Short- and long-run relationships between public sentiment and policy salience, conditional on the misery index. Lines represent predicted instantaneous effects (left) and total effects (right) on mentions in Queen's Speeches (QS) as the result of an average increase in public issue salience, holding change in the misery index at its mean value, while varying misery across its observed range. 95% confidence intervals constructed from quantiles of 5,000 samples from $\mathcal{N}_{MV}\left(\hat{\boldsymbol{\theta}}, \mathbb{C}(\hat{\boldsymbol{\theta}})\right)$ are shaded in gray.

9 Pretesting data

In this section I provide a sample of the pretesting exercises described in the text. In the Blaydes and Kayser (2011) and Morgan and Kelly (2013) applications, the data are organized into panels, and so I present just a small subset of results run separately on each time series, separated by unit (i.e., country). Similarly, I present just a few ACF and PACF plots, among the approximately 1,000 generated in analyzing these data. The replication archive contains code to reproduce all visuals and tests of interest. See Box-Steffensmeier et al. (2014) for a thorough guide on interpreting the output of these tests.

9.1 Blaydes and Kayser (2011)

Table 4: Univariate ADF tests, Blaydes and Kayser (2011) replication

Variable	Mean p-value	% tests where null rejected
Calories	0.56	0.04
ΔCalories	0.34	0.26
GDP	0.63	0.04
ΔGDP	0.33	0.16
Regime type	0.42	0.00
ΔRegime type	0.51	0.00
GDP $\times$ Regime type	0.71	0.00
Δ (GDP $\times$ Regime type)	0.28	0.57

Table 5: Tests for cointegration, Blaydes and Kayser (2011) replication

Test	Test stat.	p-value
Ljung-Box	100.86	0.00
ADF	-17.39	0.01
KPSS	0.00	0.10
Engle-Granger	-11.07	0.00

Table 6: Johansen test for cointegration, Blaydes and Kayser (2011) replication

		λ_{TRACE}		
Null hypothesis	Alt. hypothesis	Test value	95% crit. value	90% crit. value
$r = 0$	$r > 0$	574.11	53.12	49.65
$r \leq 1$	$r > 1$	372.96	34.91	32.00
$r \leq 2$	$r > 2$	211.60	19.96	17.85
$r \leq 3$	$r > 3$	87.84	9.24	7.52

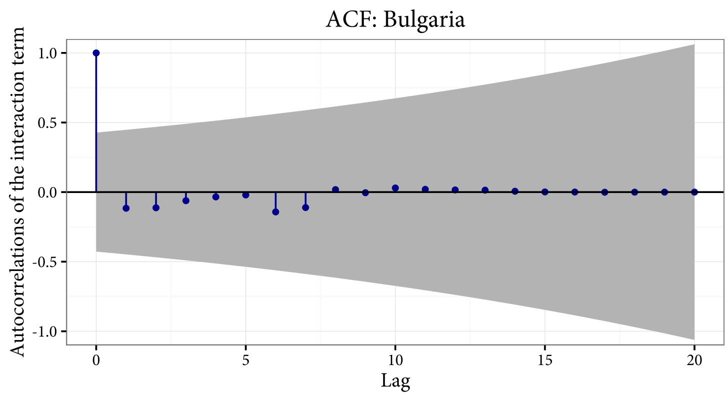

Figure 12: Sample ACF plot from the B & K application. Here the series in question is the interaction of GDP and regime type in Bulgaria.

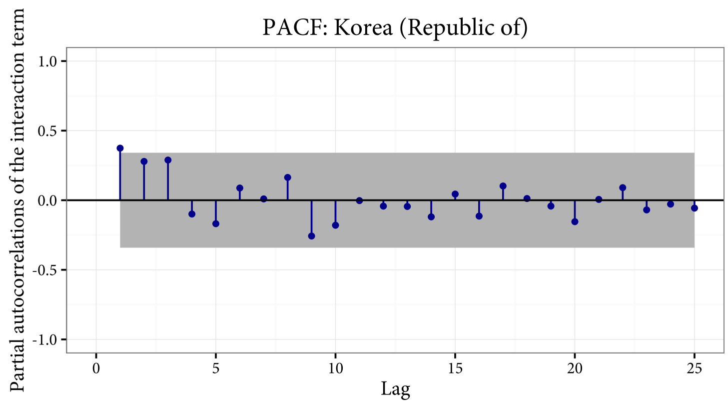

Figure 13: Sample PACF plot from the B & K application. Here the series in question is caloric consumption in South Korea.

9.2 Morgan and Kelly (2013)

Table 7: Univariate unit root tests of the interaction term, Morgan and Kelly (2013) replication

Variable	Mean p-value	% tests where null rejected
ADF test	0.55	0.04
Ljung-Box test	0.00	1.00
KPSS test	0.03	0.79
ADF test, first difference	0.34	0.21
Ljung-Box test, first difference	0.26	0.38
KPSS test, first difference	0.07	0.33

Table 8: Tests for cointegration, Morgan and Kelly (2013) replication

Test	Test stat.	p-value
Ljung-Box	27.04	0.13
ADF	−6.23	0.01
KPSS	0.09	0.10
Engle-Granger	−4.98	0.00

Table 9: Johansen test for cointegration, Morgan and Kelly (2013) replication

λ_{TRACE}				
Null hypothesis	Alt. hypothesis	Test value	95% crit. value	90% crit. value
$r = 0$	$r > 0$	321.07	202.92	196.37
$r \leq 1$	$r > 1$	237.02	165.58	159.48
$r \leq 2$	$r > 2$	173.34	131.70	126.58
$r \leq 3$	$r > 3$	130.05	102.14	97.18
$r \leq 4$	$r > 4$	90.93	76.07	71.86
$r \leq 5$	$r > 5$	61.22	53.12	49.65
$r \leq 6$	$r > 6$	37.84	34.91	32.00
$r \leq 7$	$r > 7$	20.65	19.96	17.85
$r \leq 8$	$r > 8$	6.71	9.24	7.52

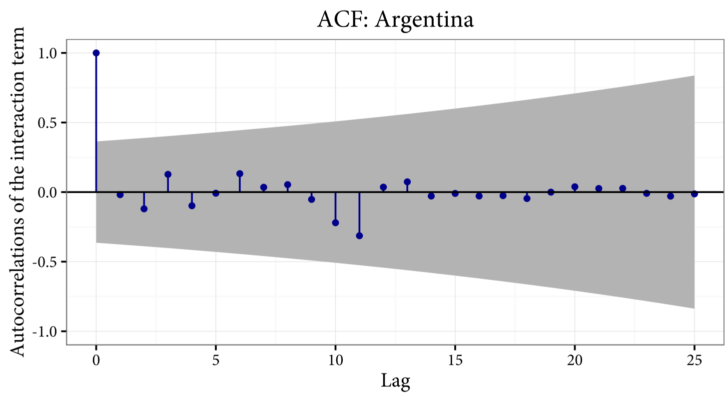

Figure 14: Sample ACF plot from the M & K application. Here the series in question is the interaction of growth and HCS in Argentina.

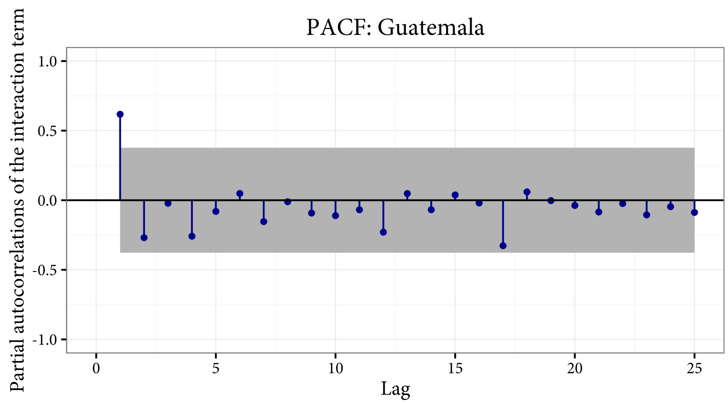

Figure 15: Sample PACF plot from the M & K application. Here the series in question is the interaction of growth and HCS in Guatemala.

9.3 Jennings and John (2009)

Table 10: p-values for univariate unit root tests by variable, Jennings and John (2009) replication

	ADF test	Ljung-Box test	KPSS test
QS	0.42	0.00	0.10
MIP	0.98	0.00	0.05
Misery	0.94	0.00	0.04
MIP $\times$ Misery	0.96	0.00	0.05
ΔQS	0.03	0.07	0.10
ΔMIP	0.01	0.34	0.04
ΔMisery	0.01	0.93	0.10
Δ MIP $\times$ Misery	0.01	0.68	0.10

Table 11: Tests for cointegration, Jennings and John (2009) replication

Test	Test stat.	p-value
Ljung-Box	37.52	0.01
ADF	−3.91	0.02
KPSS	0.31	0.10
Engle-Granger	−4.30	0.00

Table 12: Johansen test for cointegration, Jennings and John (2009) replication

λ_{TRACE}				
Null hypothesis	Alt. hypothesis	Test value	95% crit. value	90% crit. value
$r = 0$	$r > 0$	47.07	53.12	49.65
$r \leq 1$	$r > 1$	21.83	34.91	32.00
$r \leq 2$	$r > 2$	8.00	19.96	17.85
$r \leq 3$	$r > 3$	0.39	9.24	7.52

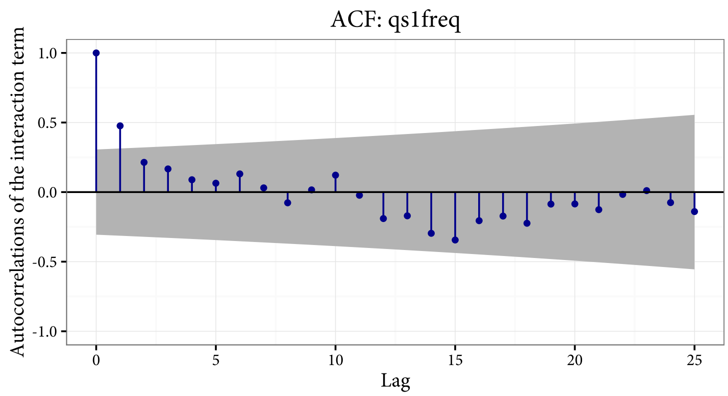

Figure 16: Sample ACF plot from the J & J application. Here the series in question is the number of economic policy mentions in the Queen's Speech.

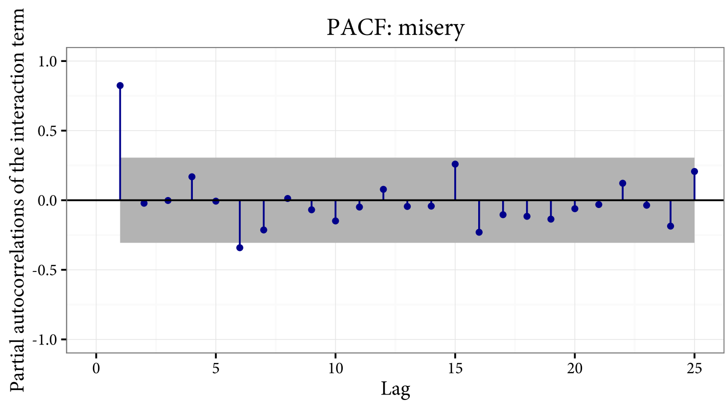

Figure 17: Sample PACF plot from the J & J application. Here the series in question is the misery index.

References

Blaydes, Lisa, and Mark Andreas Kayser. 2011. "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55 (4): 887–908.

Bohrnstedt, George W., and Arthur S. Goldberger. 1969. "On the Exact Covariance of Products of Random Variables." *Journal of the American Statistical Association* 64 (238): 1439–1442.

Box-Steffensmeier, Jan, John R. Freeman, Jon Pevehouse, and Matthew Perry Hitt. 2014. *Time Series Analysis for the Social Sciences.* New York: Cambridge University Press.

Davidson, Russell, and James G. MacKinnon. 2003. *Econometric Theory and Methods.* Oxford: Oxford University Press.

De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52 (1): 184–200.

Gallup. 2001. *Gallup Political and Economic Index.* London: Gallup Organization.

Goodman, Leo A. 1960. "On the Exact Variance of Products." *Journal of the American Statistical Association* 55 (292): 708–713.

Heston, Alan, Robert Summers, and Bettina Aten. 2001. "Penn World Table Version 6.0." Center for International Comparisons at the University of Pennsylvania (CICUP).

Jennings, Will, and Peter John. 2009. "The Dynamics of Political Attention: Public Opinion and the Queen's Speech in the United Kingdom." *American Journal of Political Science* 53 (4): 838–854.

King, Anthony, and Robert Wybrow. 2001. *British Political Opinion 1937-2000.* London: Politicos.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44 (2): 341–355.

Morgan, Jana, and Nathan J. Kelly. 2013. "Market Inequality and Redistribution in Latin America and the Caribbean." *Journal of Politics* 75 (3): 672–685.

Morley, Samuel. 2001. *The Income Distribution Problem in Latin America and the Caribbean.* Santiago, Chile: United Nations Press.

Solt, Frederick. 2009. "Standardizing the World Income Inequality Database." *Social Science Quarterly* 90 (2): 231–242.