

Appendix: Conditions under which RETRAFECTION

Estimation has a Causal Interpretation

The causal question of interest is how would the observed impact measure sequence $\{Y_t\}_t^N$ change if a development program were to enter and remain in an area beginning in year \bar{t} relative to no program access for this local population during the same sequence of time?

We define the following notation so that we can analyze RETRAFECTION estimation in a counterfactual framework.

- $t = 1, 2, \dots, N$ denotes the time period. Note that we use N here because T has been used to indicate treatment in section 2.
- D_i^t : a sequence of binary treatment history for individual i up to time t . Note that the subscript here denotes the *history* up to time t . That is, $D_i^t = (D_{i1}, D_{i2}, \dots, D_{it})$. $D_i^N = (D_{i1}, D_{i2}, \dots, D_{iN})$, or the sequence of binary treatment for the whole period.
- D_0^N : reference treatment, $D_0^N = (0, 0, \dots, 0)_{N \times 1}$. This is a particular realization of D_i^N .
- D_1^N : comparison treatment, $D_1^N = (0, 0, \dots, 1, \dots, 1)_{N \times 1}$, where 1 starts from year \bar{t} . This is another particular realization of D_i^N .
- Y_{it}^0 and Y_{it}^1 : *potential* outcomes for individual i at time period t given reference treatment and comparison treatment, respectively, where $t = 1, 2, \dots, N$. Note that at most one of the potential outcomes can be observed for any individual i .
- $D_i^{t-k, t+k}$: a sequence of binary treatment variables during the event window for t , where superscript $(t-k, t+k)$ implies the event window from period $t-k$ to $t+k$. That is, $D_i^{t-k, t+k} = (D_{it-k}, D_{it-k-1}, \dots, D_{it}, D_{it-1}, D_{it-k})$.
- \tilde{Z}_i^t : a sequence of the history of observed potential causes other than the treatment for individual i up to time t . That is, $\tilde{Z}_i^t = (\tilde{Z}_{i1}, \tilde{Z}_{i2}, \dots, \tilde{Z}_{it})$.
- Z_i^t : a sequence of the history of unobserved potential causes other than the treatment for individual i up to time t . That is, $Z_i^t = (Z_{i1}, Z_{i2}, \dots, Z_{it})$.
- W_i^t : a sequence of the history of observed proxies for individual i up to time t . That is, $W_i^t = (W_{i1}, W_{i2}, \dots, W_{it})$.
- \tilde{X}_i^t : a set of observed covariates, including observed causes and proxies. That is, $\tilde{X}_i^t \equiv (\tilde{Z}_i^t, W_i^t)$.

- $X_i^t: X_i^t \equiv D_i^{t, k, t, k}, \tilde{X}_i^t$.

Note that RETRAFECT assumes that only treatments within this event window can have effects on Y_{it} . Following the framework established in White (2006), the outcome is determined by a response function:

$$y_{it} = c_i^t(d_i^{t, k, t, k}, \tilde{z}_i^t, z_i^t), \quad (\text{A1})$$

where we assume that the outcome at time t depends on i and depends on the sequence of treatment realizations during the event window $d_i^{t, k, t, k}$, the observed causes history \tilde{z}_i^t and the unobserved causes history z_i^t . If we further assume that the outcome depends on i only through the arguments of the response function, the functional form c does not depend on i . Thus,

$$\begin{aligned} y_{it} &= c_i^t(d_i^{t, k, t, k}, \tilde{z}_i^t, z_i^t) \\ &= c^t(d_i^{t, k, t, k}, \tilde{z}_i^t, z_i^t). \end{aligned} \quad (\text{A2})$$

Given (\tilde{z}^t, z^t) , the effect of the event at time t is

$$\begin{aligned} \Delta_{it}(\tilde{z}^t, z^t) &= c^t(d_i^{t, k, t, k}, \tilde{z}_i^t, z_i^t) - c^t(d_i^{t, k, t, k}, \tilde{z}_i^t, z_i^t) \\ &\equiv c_1^t(\tilde{z}_i^t, z_i^t) - c^t(\tilde{z}_i^t, z_i^t). \end{aligned}$$

Write

$$\begin{aligned} Y_{it} &\equiv c^t(\tilde{Z}_i^t, Z_i^t), \text{ and} \\ Y_{it}^1 &\equiv c_1^t(\tilde{Z}_i^t, Z_i^t). \end{aligned} \quad (\text{A3})$$

At period t , the expected effect for treated for given x_i^t is

$$\begin{aligned} \Delta_{1it}(x_i^t) &\equiv E[Y_{it}^1 | D_i^{t, k, t, k} = 1, \tilde{X}_i^t = \tilde{x}_i^t] \\ &\quad - E[Y_{it} | D_i^{t, k, t, k} = 1, \tilde{X}_i^t = \tilde{x}_i^t] \\ &\equiv \tilde{\mu}_{1t}(x_i^t) - \tilde{\mu}_{1t}(x_i^t), \end{aligned} \quad (\text{A4})$$

where

$$\tilde{X}_i^t \equiv \tilde{Z}_i^t, W_i^t, X_i^t \equiv D_i^{t, k, t, k}, \tilde{X}_i^t$$

and $\{W_{it}\}$ are observable proxies. Note that \tilde{X}_i^t includes observed causes and proxies.

For the estimation of $\tilde{\mu}_{1t}$, assuming the unconfoundedness condition

$$\begin{aligned} Z_i^t &\perp D_i^{t, k, t, k} | \tilde{X}_i^t, \text{ then} \\ \tilde{\mu}_{1t}(x_i^t) &= \tilde{\mu}_{1t}(x_i^t) \equiv E[Y_{it} | D_i^{t, k, t, k} = 1, \tilde{X}_i^t = \tilde{x}_i^t]. \end{aligned} \quad (\text{A5})$$

Assuming only the contemporary values \tilde{Z}_{it} and Z_{it} affect the outcome, and the true relationship is a linear parametric one, i.e.,

$$\begin{aligned}
Y_{it} &= c^t D_i^{t, k, t, k}, \tilde{Z}_i^t, Z_i^t \\
&= \sum_{t, \bar{t}, k} \tau_t D_{it} + \tilde{Z}'_{it} \theta + Z_{it},
\end{aligned} \tag{A6}$$

Where \tilde{Z}'_{it} includes a constant term. In addition assume

$$Z_{it} \perp D_i^{t, k, t, k} | \tilde{X}_{it},$$

then

$$\begin{aligned}
&E[Y_{it} | D_i^{t, k, t, k}, \tilde{X}_i^t] \\
&= \tilde{Z}'_{it} \theta + \sum_{t, \bar{t}, k} \tau_t D_{it} + E[Z_{it} | D_i^{t, k, t, k}, \tilde{X}_i^t] \\
&= \tilde{Z}'_{it} \theta + \sum_{t, \bar{t}, k} \tau_t D_{it} + E[Z_{it} | \tilde{X}_i^t] \\
&= \tilde{Z}'_{it} \theta + \sum_{t, \bar{t}, k} \tau_t D_{it} + \tilde{Z}'_{it} \gamma_1 + W'_{it} \gamma_2 \\
&= \tilde{X}'_{it} \beta + \sum_{t, \bar{t}, k} \tau_t D_{it}.
\end{aligned} \tag{A7}$$

Thus, τ_t can be consistently estimated by RETRAFECT.

Note that

$$Y_{it}^1 = \sum_{t, \bar{t}, k} \tau_t D_{it} + \tilde{Z}'_{it} \theta + Z_{it}$$

and

$$Y_{it} = \sum_{t, \bar{t}, k} \tau_t D_{it} + \tilde{Z}'_{it} \theta + Z_{it}.$$

So the causal effect at time t would be

$$Y_{it}^1 - Y_{it} = \sum_{t, \bar{t}, k} \tau_t (D_{it} - D_{it}). \tag{A8}$$

Because the reference treatment and comparison treatment are known, and because we can estimate τ_t , we can estimate the causal effect.

In the setting of credit availability, T_{it} is the treatment D_{it} . The effect lasts for up to k periods after the initial availability. The effect from the treatment at $t + 1$ to $t + k$ should be viewed as the effect from the expected treatment which is known at time t . The effect of expected availability started from at most k years before. The effect does not depend on which year is the initial year when microfinance institution is available.

References

- Amendariz, B. and Morduch, J. (2005). Measuring impacts. *Economics of microfinance* (pp. 207-229). The MIT Press: Cambridge, Massachusetts.
- Attanasio, Orazio, Britta Augsborg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart (2011) Group Lending or Individual Lending? Evidence from a Randomised Field Experiment in Mongolia, University of College London working paper.
- Banerjee, A., Duflo, E., Glennerster, R. and Kinnan, C. (2009). The miracle of microfinance? Evidence from randomized evaluation. *Centre for Micro Finance, IFMR Research, Working Paper Series No. 31*.
- Berhane, G. and Gardebroke, C. (2011). Does microfinance reduce rural poverty? evidence based on household panel data from northern Ethiopia. *American Journal of Agricultural Economics, forthcoming*.
- Bhalotra, S., Copestake, J., and Johnson, S. (2001). Assessing the Impact of Microcredit: A Zambian Study. *The Journal of Development Studies*, 37(4), 81-100. CGAP (2002). Microfinance and the millennium development goals, CGAP donor brief No. 9, December 2002, CGAP, http://www.cgap.org/gm/document-1.9.2422/DonorBrief_09.pdf.
- CGAP (2003). Helping to improve donor effectiveness: the Impact of Microfinance. CGAP donor brief No. 13.
- Chamberlain, G. (1979). Analysis of covariance with qualitative data. *National Bureau of Economic Research, Working Paper No. 325*.
- Coleman, B. (2002). Microfinance in northeast Thailand: Who benefits and how much? ADB, ERD working paper series no. 9.
- Crépon, Bruno, Florencia Devoto, Esther Duflo and William Parienté (2011). "Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation" MIT/JPAL working paper, March.
- Dingcong, C., Infantado, C., Kondo, T. and Orbeta Jr, A., (2008). Impact of microfinance on rural households in the Philippines. *Philippines Institute for Development Studies*. Discussion Paper series no. 2008.05.
- Dizon, F. (2009). Rural out-migration and farm asset accumulation: the case of Guatemala. (Master's thesis and presented at the Pacific Development Conference, March 2009)

Table 2C: FE OLS RETRAFACT Regression on Consumer Durables Purchase

VARIABLES	--- Credit Availability Window ---					Uptake
	New television	New bicycle	New cellphone	New stove	Any Consumer Durable	Any Consumer Durable
4 years before	0.008 (0.010)	-0.002 (0.011)	-0.001 (0.005)	0.002 (0.002)	0.008 (0.014)	0.004 (0.011)
3 years before	0.004 (0.013)	-0.010 (0.013)	-0.002 (0.008)	0.000 (0.002)	0.003 (0.017)	0.025** (0.012)
2 years before	0.044*** (0.015)	0.004 (0.017)	-0.000 (0.011)	0.006 (0.004)	0.040* (0.022)	0.023 (0.014)
1 year before	0.034** (0.017)	0.040** (0.019)	0.002 (0.014)	0.005 (0.005)	0.079*** (0.025)	0.031** (0.014)
Credit access yr	0.015 (0.021)	0.002 (0.023)	-0.009 (0.018)	0.003 (0.007)	0.002 (0.032)	0.092*** (0.031)
1 year after	0.005 (0.023)	0.014 (0.025)	0.024 (0.023)	0.011 (0.008)	0.040 (0.034)	0.088*** (0.027)
2 years after	0.001 (0.027)	0.072** (0.029)	0.059** (0.026)	0.017** (0.007)	0.115*** (0.041)	0.103*** (0.032)
3 years after	-0.024 (0.024)	0.029 (0.027)	0.016 (0.025)	0.002 (0.007)	0.020 (0.036)	0.092** (0.037)
4 years after	-0.029 (0.038)	0.050 (0.043)	0.015 (0.047)	-0.012 (0.021)	0.006 (0.063)	0.103 (0.067)
education	0.019*** (0.004)	0.005 (0.005)	0.021*** (0.005)	0.004* (0.002)	0.034*** (0.007)	0.029*** (0.007)
age	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.003 (0.002)	-0.004** (0.002)
age-squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)
retail trade	-0.003 (0.006)	-0.006 (0.007)	-0.003 (0.006)	0.001 (0.002)	-0.009 (0.010)	-0.006 (0.012)
agric*retail	0.010 (0.008)	0.016* (0.009)	0.007 (0.008)	0.002 (0.004)	0.019 (0.013)	0.015 (0.013)
Constant	0.007 (0.027)	0.092*** (0.035)	-0.005 (0.026)	-0.008 (0.012)	0.086* (0.045)	0.140*** (0.031)
Observations	7,020	7,020	7,020	7,020	7,020	7,020
R-squared	0.031	0.043	0.093	0.019	0.113	0.013
<i>F</i> -stat: post- vs. pre-treatment	2.98*	2.43	2.02	0.03	0.18	5.63**
<i>p</i> -value	0.085	0.119	0.156	0.853	0.671	0.025

Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

**Table 3: ITT, Take-Up, and LATE Estimations
Enterprise Investment, Housing Improvements, and Consumer Durables**

VARIABLES	New machines	New lands	New livestock	New store	New investments
Microfinance ITT	-0.002 (0.007)	-0.000 (0.008)	0.022* (0.012)	0.008** (0.004)	0.023* (0.013)
Microfinance Take-up	0.009*** (0.003)	0.031*** (0.009)	0.080*** (0.012)	0.019*** (0.004)	0.127*** (0.013)
Microfinance LATE	-0.002 (0.027)	-0.008 (0.037)	0.127* (0.066)	0.050* (0.027)	0.181** (0.076)
VARIABLES	New roof	New wall	New floor	New toilet	All Home Improvements
Microfinance ITT	0.007 (0.010)	-0.000 (0.005)	-0.010 (0.008)	0.006 (0.008)	0.009 (0.016)
Microfinance Take-up	0.039*** (0.010)	0.001 (0.003)	0.008* (0.004)	0.021*** (0.008)	0.063*** (0.015)
Microfinance LATE	0.063 (0.055)	0.001 (0.022)	-0.031 (0.027)	0.052 (0.037)	0.100 (0.066)
VARIABLES	New televisions	New bicycles	New cell phone	New stove	All Consumer Durables
Microfinance ITT	-0.034*** (0.008)	-0.016 (0.019)	0.006 (0.015)	0.002 (0.003)	-0.044*** (0.011)
Microfinance Take-up	0.016* (0.008)	0.042*** (0.015)	0.059*** (0.015)	0.009 (0.005)	0.089*** (0.022)
Microfinance LATE	-0.163** (0.065)	-0.092 (0.076)	0.034 (0.067)	0.012 (0.025)	-0.194* (0.104)

Each cell gives a single impact coefficient. Credit availability instruments for take-up in LATE estimations. Clustered-robust standard errors in parentheses with region-year fixed effects made up of paired villages within a region and a single year. Regressions include controls for education age, age², and type of enterprise.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.