Extrapolation of Treatment Effect Estimates Across Contexts and Policies: An Application to Cash Transfer Experiments

Kensuke Maeba

Northwestern University

January, 2023

Two extrapolation approaches

- Evidence-based policymaking is challenging
  - External validity not always guaranteed
  - From which evidence to extrapolate?

1. Extrapolation across contexts (within policies)
   - Policy A in context X $\Leftarrow$ Policy A in **context Y**
2. Extrapolation across policies (within contexts)
   - Policy A in context X $\Leftarrow$ **Policy B** in context X

- **Do two extrapolations result in different predictions?**
  - No theory & limited empirical evidence
  - Comparative analysis using **cash transfer experiments**

Cash transfer experiments

- Conditional cash transfers (CCTs) on school enrollment rates
  - Common version: conditional on regular school attendance
  - Popular anti-poverty policy (e.g. PROGRESA, Bolsa Familia, RPS)
  - Abundant empirical evidence showing **positive effects**
- Varying cost-effectiveness across programs
  - Heterogeneity in effect size Effect size by transfer amount
- Data: two RCTs in Malawi & Morocco
  - CCTs implemented in both RCTs
  - "Unconditional" cash transfers in Moroccan RCT

Overview

$$\text{Malawi CCTs} \quad \underbrace{\Rightarrow}_{\text{Across contexts}} \quad \text{Morocco CCTs} \quad \underbrace{\Leftarrow}_{\text{Across policies}} \quad \text{Morocco LCTs}$$

- Predict average effect of Moroccan CCTs on school enrollment rates
  - CCTs: cash transfers if school attendance $>$ threshold
  - LCTs: cash transfers $\perp\!\!\!\perp$ school attendance
- Compare predictions with actual treatment effect estimate
  - Predictions via **standard** method (Data-driven vs Structural)
  - Predictions via **same** method (Structural vs Structural)
- Identify source of prediction differences through structural model

Preview of findings

- Moroccan LCTs make more accurate predictions than Malawi CCTs
  - Estimated effect = 5.7 ppt: 5.9 ppt vs 21.2 ppt
  - Statistically significant difference only for across-contexts
  - Relative performance unchanged when using same method
- **(Perceived) returns to schooling** explain prediction differences
  - Key parameters in schooling decisions: cost today vs return tomorrow
  - Estimated cost of schooling: similar for same policy
  - Estimated returns to schooling: similar for same context
  - Discussion on varying perceived returns to schooling across contexts

Contributions to literature

- **Empirical investigation of out-of-sample predictions**
  - Pritchett and Sandefur (2015); Gechter et al. (2018)
  - Prediction performance of across-policies extrapolation
  - Comparative analysis of two extrapolation approaches
- Structural estimation with RCTs in development economics
  - Todd and Wolpin (2006); Attanasio et al. (2012)
  - Identification of flexible model about schooling
- New identification of dynamic discrete choice model
  - Scott (2014); Kalouptsidi et al. (2021)
  - No rational expectations assumption
  - First application to schooling decisions

Two cash transfer experiments

- Malawi (Baird et al., 2011)
  - Treatment: CCTs
  - Target: **girls at secondary school ages** (13 - 22 years old)
  - Teenage pregnancy & marriage as main driver of dropout

- Morocco (Benhassine et al., 2015)
  - Treatment: CCTs and LCTs (Labeled Cash Transfers)
  - Target: **boys and girls at primary school ages** (6 - 15 years old)
  - Various reasons to drop out: school quality, financially, domestic work
  - LCTs $\neq$ UCTs due to **endorsement effects**

▸ Summary statistics

Treatment effect estimates on school enrollment rates

	Malawi	Morocco	
	CCTs	CCTs	LCTs
Treatment	0.0369*	0.0567***	0.0726***
	(0.0200)	(0.0106)	(0.0107)
Control mean	0.896***	0.894***	0.893***
	(0.0154)	(0.00951)	(0.00833)
Obs.	1490	4982	3018

*** p<0.01 ** p<0.05 * p<0.1

- Estimated effects greater (not statistically)
  - for LCTs within contexts
  - for Morocco within policies

▸ Why LCTs more effective? ▸ Back

Across-context: linear projection

- Heterogeneous treatment effect approach
  1. Estimate linear regression with Malawi CCTs data

$$d_i = W_i' \beta^{\text{HTE}} + \beta_0^{\text{HTE}} \text{Treatment}_i + \sum_{k=1}^{K} \gamma_k^{\text{HTE}} \text{Treatment}_i \times w_{ik} + \omega_i$$

  2. Predict d_i in Moroccan CCTs data using estimated linear model
  3. Estimate treatment effect via OLS

$$\hat{d}_i = \alpha_0^{\text{HTE}} + \alpha_1^{\text{HTE}} \text{Treatment}_i + \text{Stratum}_i + \nu_i^{\text{HTE}}$$

- W_i = (age, education, per-capita income, school costs, transfer amount)

- Key assumption: **cond. on W_i, potential outcomes $\perp\!\!\!\perp$ contexts**

▸ Propensity score weighting

Across-policies: dynamic model of schooling

- Child i makes binary schooling decisions over finite time horizon
  - Schooling ($d_i = 1$): pay school costs (s_i), add 1 year of education
  - Non-schooling ($d_i = 0$): consume (per-capita) income (y_i)
  - Flow utility function: $u(c) = \theta \ln(c_i)$

- At terminal period T, child receives lump-sum returns
  - Returns are a function of education: $R_i = R(e_{i,T}) = R\left(e_{i,t} + \sum_{\tau=t}^{T-1} d_{i,\tau}\right)$

- Cash transfers (z_i) relax budget constraints exogenously
  - Budget constraint: $c_i = y_i - d_i \times s_i$
  - CCTs lower school costs: $s_i = s_i - z_i$ if i in treatment group
  - LCTs increase income: $y_i = y_i + z_i$ if i in treatment group

Formal representation

Prediction with θ & $R\left(e_{i,T}\right)$

- Predict prob. of schooling under Moroccan CCTs

$$\hat{P}\left(d=1|e_{i,2},y_{i,2},s_{i,2}\right)=\frac{\exp\left(\hat{\theta}\ln\left(\frac{y_{i,2}-s_{i,2}}{y_{i,2}}\right)+\beta\widehat{\Delta R}\left(e_{i,2}\right)\right)}{1+\exp\left(\hat{\theta}\ln\left(\frac{y_{i,2}-s_{i,2}}{y_{i,2}}\right)+\beta\widehat{\Delta R}\left(e_{i,2}\right)\right)}$$

- Estimate treatment effect via OLS

$$\hat{P}\left(d=1|e_{i,2},y_{i,2},s_{i,2}\right)=\delta_1+\delta_2\text{Treatment}_i+\text{Stratum}_i+\nu_i$$

Structural estimation: sketch

1. Compute probability of schooling in all possible state values
   - Frequency estimator or **parametric to smooth across states**

2. Consider two hypothetical paths for each individual
   - (Baseline, Midline) = (Schooling, Non-schooling), (Non-schooling, Schooling)
   - Same years of education in next period after Midline
   - **Continuation values after Midline fixed**

3. Compare changes in probability of schooling across these paths
   - Differences only in flow utilities between Baseline and Midline
   - **Δ in prob $\Leftrightarrow$ Δ in flow utility, shifted by cash transfers**
   - Size of Δ in flow utility different across CCTs and LCTs

▸ Formal derivation

Identification of θ using cash transfers

$$\underbrace{\ln \frac{P_{it}^{1}}{1-P_{it}^{1}} - \beta \ln \frac{P_{i,t+1}^{2}}{1-P_{i,t+1}^{3}}}_{\text{Change in prob. of schooling across paths}} = {\color{red}\theta} \underbrace{\left\{ \ln \left(\frac{y_{it}-s_{it}}{y_{it}} \right) - \beta \ln \left(\frac{y_{i,t+1}-s_{i,t+1}}{y_{i,t+1}} \right) \right\}}_{\text{Change in flow utility of schooling across paths}}$$

$$+ \underbrace{\beta \left(\eta_{it}(1) - \eta_{it}(0) \right)}_{\text{Expectation errors at baseline (Unobservable)}}$$

- Endogeneity: **expectation errors correlated with income levels**
  - e.g.) information friction, ability to forecast

- **Treatment assignment** as an IV
  - Relevance: cash transfers (only) shift school costs (↓) or income (↑)
  - Exclusion: assignment $\perp\!\!\!\perp$ baseline expectation

Identification of $R(e;x)$

- Identification of terminal payoffs depends on time horizon

- Set terminal period at one period after Midline ($T = 3$)

- Expand prob. of schooling at Midline ($t = 2$)

$$\underbrace{\ln \frac{P^1_{i,2}}{1 - P^1_{i,2}}}_{\text{Odds of schooling}} = \underbrace{\theta \ln \left(\frac{y_{i,2} - s_{i,2}}{y_{i,2}} \right)}_{\text{Flow utility of schooling}} + \beta \underbrace{\left(R(e_{i,2} + 1; x_{i,2}) - R(e_{i,2}; x_{i,2}) \right)}_{\equiv \Delta R(e_{i,2}; x_{i,2})}$$

- $\Delta R(e_{i,2}; x_{i,2})$: Perceived relative returns to schooling
  - **Schooling unexplained by contemporaneous effects of cash transfers**
  - Consequences of schooling decisions after RCTs
  - Heterogenous across state values

Across-contexts (reduced-form) vs Across-policies (structural)

	Target	Across-contexts	Across-policies
Treatment	0.0567***	0.212***	0.0590***
	(0.0106)	(0.00442)	(0.00545)
Control mean	0.894***	1.127***	0.941***
	(0.00951)	(0.00373)	(0.00531)
Obs.	4982	4982	4982
= Target TE		0.000	0.674
= Target control mean		0.000	0.000

Note: *** $p<0.01$ ** $p<0.05$ * $p<0.1$

- Across-contexts: significantly overpredict
- Across-policies: numerically and statistically accurate

▸ Linear vs Reweighting

Across-contexts (structural) vs Across-policies (structural)

	Target	Across-contexts	Across-policies
Treatment	0.0567***	0.0431***	0.0590***
	(0.0106)	(0.00465)	(0.00545)
Control mean	0.894***	0.702***	0.941***
	(0.00951)	(0.00390)	(0.00531)
Obs.	4982	4982	4982
= Target TE		0.004	0.674
= Target control mean		0.000	0.000

Note: *** p<0.01 ** p<0.05 * p<0.1

- Across-contexts: slightly but statistically significantly underpredict

How to extrapolate $\Delta R(e; x)$ Why improvement via structural method?

Comparison of θ and $\Delta R(e; x)$

- Comparative analysis at model parameter level
  - Structural model estimated for each intervention (including Moroccan CCTs)
  - Comparison of parameter estimates across interventions
  - Identification of varying DGP if model is true Check model fit

- Two parameters that determine schooling decisions
  - Flow utility cost of schooling: $\theta \ln(y - d \times s)$
  - Perceived relative returns to schooling: $\Delta R(e; x)$

Comparison of θ

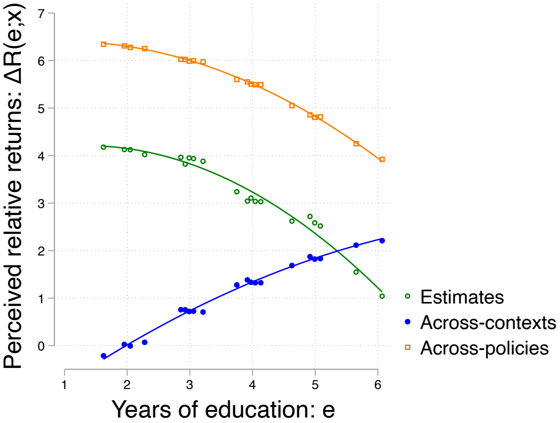
	Malawi	Morocco	
	CCTs	CCTs	LCTs
θ	1.008***	2.670***	38.90***
	(0.256)	(0.454)	(11.30)
Obs.	1479	4981	3016
1st stage F statistics	113.011	3843.510	25.483
= target θ	0.000		0.001

Note: I report the Kleiberge-Paap F statistics for weak identification. *** p<0.01 ** p<0.05 * p<0.1

- Numerically closer estimates if same policy

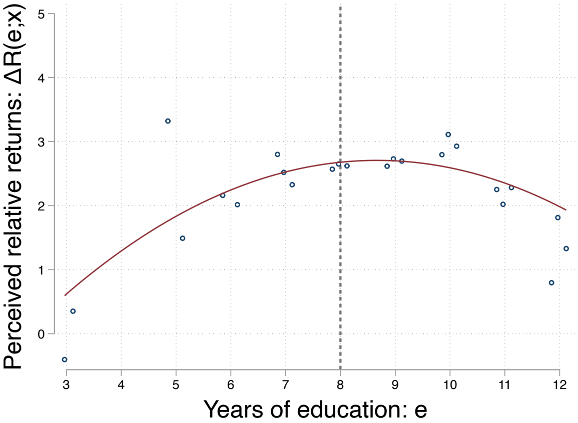
Compare elasticity
Replace with estimated values

Comparison of $\Delta R(e;x)$



- Estimate of $\Delta R(e;x)$: downward sloping due to increasing opportunity costs
- $\Delta R(e;x)$ from Moroccan LCTs: parallel to target values
- $\Delta R(e;x)$ from Malawi CCTs: opposite direction

Why extrapolation of $\Delta R(e;x)$ from Malawi CCTs wrong?



- Increase during primary education
- Start decreasing at which school dropout becomes more realistic
- Naive extrapolation: use increasing part to predict in target context

What explains variation in perceived returns across contexts?

- Perceived relative returns to schooling are key for prediction accuracy
  - Negative correlation bet. estimates and extrapolation across contexts
- To what extent differences in $\Delta R(e;x)$ explained by observables?
  1. Normalize years of education of Malawi sample
  2. Restrict sample by age and sex

Normalization of years of education

	Target	Across-contexts		
Treatment	0.0567***	0.0431***	0.0150***	0.0147***
	(0.0106)	(0.00465)	(0.000788)	(0.00265)
Control mean	0.894***	0.702***	0.903***	0.875***
	(0.00951)	(0.00390)	(0.000674)	(0.00221)
Obs.	4982	4982	4982	4982
Normalization			✓	✓
Sample restriction				✓
= Target TE		0.004	0.000	0.000
= Target control mean		0.000	0.000	0.000

Note: *** p<0.01 ** p<0.05 * p<0.1

- Extrapolation using decreasing part of $\Delta R(e;x)$
- Treatment effect still underpredicted (while overall prediction improved)

▸ RF with normalization

Heterogeneity by sex of Moroccan sample

	Target		Across-contexts	
	Boys	Girls	Boys	Girls
Treatment	0.0479***	0.0681***	0.0151***	0.0148***
	(0.0130)	(0.0144)	(0.00109)	(0.00129)
Control mean	0.912***	0.871***	0.902***	0.903***
	(0.0115)	(0.0129)	(0.000913)	(0.00112)
Obs.	2666	2313	2666	2313
= Target TE			0.000	0.000
= Target control mean			0.000	0.000

Note: *** p<0.01 ** p<0.05 * p<0.1

- Prediction separately by sex of Moroccan sample
- No additional improvement

Heterogeneity by age for Malawi Sample

	Target		Across-contexts	
	Boys	Girls	Boys	Girls
Treatment	0.0479***	0.0681***	0.00376***	0.00402***
	(0.0130)	(0.0144)	(0.000254)	(0.000318)
Control mean	0.912***	0.871***	0.980***	0.980***
	(0.0115)	(0.0129)	(0.000211)	(0.000270)
Obs.	2666	2313	2666	2313
= Target TE			0.000	0.000
= Target control mean			0.000	0.000

Note: *** p<0.01 ** p<0.05 * p<0.1

- Prediction based on young Malawi sample (age < 17)
- No additional improvement

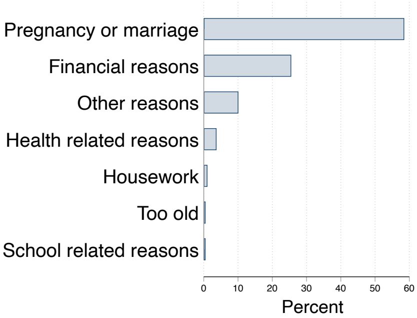
Plausible explanations for no improvement

1. Two Moroccan cash transfers are more alike than two CCTs ATE estimates
   - **Endorsement effects** Within-intervention variation
   - Confusion about conditionality of Moroccan CCTs Under perfect knowledge

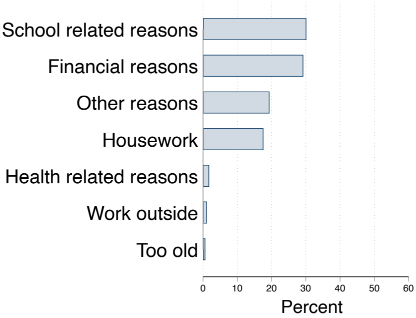
2. **Outside options are different across contexts**
   - $\Delta R(e; x)$ represents returns relative to outside options
   - Suggestive evidence by looking at primary reasons for school dropout

Varying outside options across contexts

Malawi CCTs



Moroccan CCTs



- Primary reasons for dropout differ across contexts
- Differential consequences of non-schooling

Conclusion

- When across-contexts extrapolation does not work, what can we do?
  - Status-quo approach in out-of-sample predictions
  - Empirical evidence on limitation of across-contexts extrapolation
- This paper sheds light on potentials of across-policies extrapolation
  - Similar policy that resembles how target policy works
  - Cash transfers w/o regular school attendance for CCTs
  - Proof-of-concept analysis but relevant beyond CCTs
- Across-policies dominates across-contexts due to perceived returns to schooling
  - Perceived returns to schooling more context-dependent
  - Suggestive explanations: endorsement effects & outside options

Extensions

1. Can we generalize empirical findings in this paper?
   - Three ways to generalize: contexts, policies, and methods
   - Conditions for better predictions via across-policies extrapolation

2. What features are must-have for accurate predictions?
   - WTP for precise estimates of policy effects (e.g. Hjort et al. 2021)
   - Policy designs useful for future predictions
   - Any unintended consequences of such designs?

3. Can we aggregate evidence from multiple policies?
   - Recent development on aggregation for same policies (Meager, 2019, 2022)
   - e.g.) Better predictions by using Malawi CCTs and Moroccan LCTs?

Thank you very much! Feel free to email your comments!

`kensukemaeba2022@u.northwestern.edu`

References I

Attanasio, Orazio P., Costas Meghir, and Ana Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *The Review of Economic Studies* 79 (1): 37–66. 10.1093/restud/rdr015.

Baird, Sarah, Francisco HG Ferreira, Berk Özler, and Michael Woolcock. 2014. "Conditional, Unconditional and Everything in between: A Systematic Review of the Effects of Cash Transfer Programmes on Schooling Outcomes." *Journal of Development Effectiveness* 6 (1): 1–43.

Baird, Sarah, Craig McIntosh, and Berk Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *The Quarterly journal of economics* 126 (4): 1709–1753.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. 2015. "Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education." *American Economic Journal: Economic Policy* 7 (3): 86–125. 10.1257/pol.20130225.

References II

Gechter, Michael, Cyrus Samii, Rajeev Dehejia, and Cristian Pop-Eleches. 2018. "Evaluating Ex Ante Counterfactual Predictions Using Ex Post Causal Inference." *arXiv preprint arXiv:1806.07016*.

Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini. 2021. "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities." *American Economic Review* 111 (5): 1442–1480. 10.1257/aer.20190830.

Kalouptsidi, Myrto, Paul T. Scott, and Eduardo Souza-Rodrigues. 2021. "Linear IV Regression Estimators for Structural Dynamic Discrete Choice Models." *Journal of Econometrics* 222 (1): 778–804. 10.1016/j.jeconom.2020.03.016.

Meager, Rachael. 2019. "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments." *American Economic Journal: Applied Economics* 11 (1): 57–91. 10.1257/app.20170299.

References III

Meager, Rachael. 2022. “Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature.” *American Economic Review* 112 (6): 1818–1847. 10.1257/aer.20181811.

Pritchett, Lant, and Justin Sandefur. 2015. “Learning from Experiments When Context Matters.” *American Economic Review* 105 (5): 471–475. 10.1257/aer.p20151016.

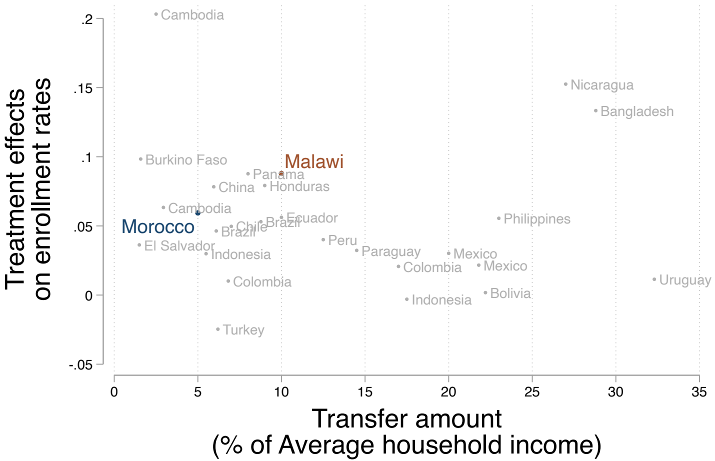
Scott, Paul. 2014. “Dynamic Discrete Choice Estimation of Agricultural Land Use.”

Stuart, Elizabeth A., Stephen R. Cole, Catherine P. Bradshaw, and Philip J. Leaf. 2011. “The Use of Propensity Scores to Assess the Generalizability of Results from Randomized Trials: Use of Propensity Scores to Assess Generalizability.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 174 (2): 369–386. 10.1111/j.1467-985X.2010.00673.x.

References IV

Todd, Petra E, and Kenneth I Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review* 96 (5): 1384–1417. 10.1257/aer.96.5.1384.

Collection of CCTs effect estimates



Source: Baird et al. (2014)

Back

Summary statistics at baseline

	Malawi		Morocco		
	Control	CCTs	Control	CCTs	LCTs
= 1 if enrollment	1.000	1.000	0.909	0.921	0.920*
Years of education	8.046	7.960	2.755	2.776*	2.764
Per-capita income (in 100 USD)	1.173	1.571	5.368	5.335	5.345
School costs (in 100 USD)	0.123	0.124	0.213	0.212	0.212
Cash transfers (in 100 USD)	NA	1.006	NA	1.054	1.057
=1 if girls	1.000	1.000	0.448	0.471*	0.486**
Age	14.964	14.740	9.889	9.910	9.912
Obs.	1145	412	1276	3706	1740
Joint F-test		0.153		0.250	0.106

Back

Discussion on LCTs vs CCTs in Morocco

- Why LCTs and CCTs comparable?
  1. Endorsement effect driving cash transfer effects
  2. CCTs misperceived as unconditional

- Why LCTs more effective than CCTs?
  - LCTs compliers: children not confident about regular school attendance

- Do Moroccan CCTs differ from Malawi CCTs?
  - Endorsement effect attached to Malawi CCTs
  - Conditionality correctly understood in Malawi CCTs

Back

Across-context: Reweighting

- Propensity score weighting (Stuart et al., 2011)
  1. Pool Malawi and Moroccan data
  2. Estimate propensity score of being in Malawi data via logit

$$\mathbf{1}\{i \in \text{Malawi CCTs}\} = W_i'\beta^{\text{PSW}} + \beta_0^{\text{PSW}}\text{Treatment}_i + u_i$$

  3. Estimate ATE with Malawi data reweighted by inverse of propensity score

$$d_i = \alpha_0^{\text{PSW}} + \alpha_1^{\text{PSW}}\text{Treatment}_i + \text{Stratum}_i + \nu_i^{\text{PSW}}.$$

- W_i = (age, education, per-capita income, school costs, transfer amount)

- Key assumption: **cond. on W_i, potential outcomes $\perp\!\!\!\perp$ contexts**

▸ Back

$$\max_{\{d_{i\tau}\}_{\tau=t}^{T-1}} E\left[\sum_{\tau=t}^{T-1} \beta^{\tau-t}\left\{\theta \ln\left(c_{i\tau}\right)+\varepsilon_{i\tau}\left(d_{i\tau}\right)\right\}+\beta^{T-t} R\left(e_{i,T}; x_{i,T}\right) \middle| \Omega_{i\tau}\right]$$

$$\text{s.t. } c_{i\tau} = y_{i\tau} - d_{i\tau} s_{i\tau}$$

$$e_{i\tau} = e_{i,\tau-1} + d_{i,\tau-1}$$

- State variables How to construct
  - e_{it} : Years of education
  - y_{it} : Per-capita income
  - s_{it} : School costs
  - ε_{it} : Preference shocks
  - $\Omega = \{e, y, s, \varepsilon\}$

- Parametric assumptions
  - Discount factor: $\beta = 0.95$
  - Preference shocks: $\varepsilon \sim$ Type 1 extreme value, i.i.d across (t, i, d)

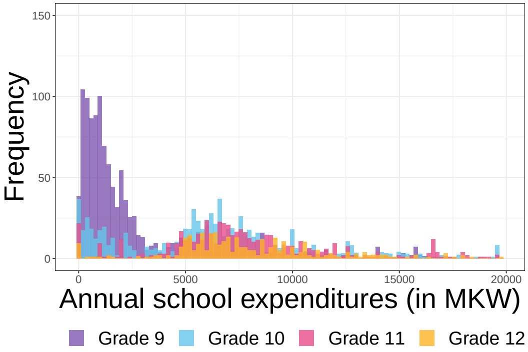
Back

Variables construction

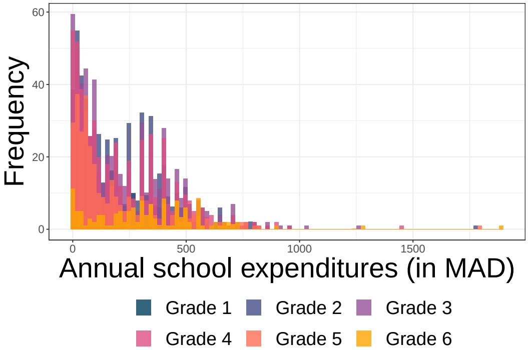
- Directly observe in data: (e_{it}, d_{it}, z_{it})
- Per-capita income: y_{it}
  - Income is unobserved or measured with errors
  - Use annual household expenditures
  - Divide by household size adjusted by OECD equivalence scale
  - If in LCTs, add cash transfer amount
- School costs: s_{it}
  - School costs that parents have to pay upfront
  - Use annual expenditures on school related stuff
  - Take median for schooling children in control group for each grade
  - If in CCTs, subtract cash transfer amount

Distribution of school costs

Malawi: secondary school



Morocco: primary school



Structural estimation: HM inversion

- Rewrite choice prob. at baseline using HM inversion ▸ How to estimate CCP

$$\underbrace{\ln \frac{P\left(d=1|e_{it},x_{it}\right)}{P\left(d=0|e_{it},x_{it}\right)}}_{\text{Odds of schooling}} = \underbrace{v\left(e_{it},x_{it},1\right)-v\left(e_{it},x_{it},0\right)}_{\text{Differences in conditional value functions}}$$

- Expand continuation values at baseline (Scott, 2014; Kalouptsidi et al., 2021)

$$\begin{aligned} v\left(e_{it},x_{it},d\right) &= \theta \ln\left(y_{it}-d\times s_{it}\right)+\beta E_x\left[\overline{V}\left(e_{i,t+1},x_{i,t+1}:\theta\right)|e_{it},x_{it},d\right] \\ &= \underbrace{\theta \ln\left(y_{it}-d\times s_{it}\right)}_{\text{Flow utility}}+\beta\left(\underbrace{\overline{V}\left(e_{i,t+1},x_{i,t+1}:\theta\right)}_{\text{Realized value function}}+\underbrace{\eta_{it}\left(d\right)}_{\text{Expectation errors}}\right) \end{aligned}$$

Structural estimation: finite dependence

- Specify decisions after intervention to have $e_{i,t+2} = e_{i,t} + 1$ for everyone

$$\overline{V}\left(e_{i,t+1}, x_{i,t+1} : \theta\right) = \begin{cases} v\left(e_{it}+1, x_{i,t+1}, 0\right) + \gamma - \ln P\left(d=0 | e_{it}+1, x_{i,t+1}\right) & \text{if } d_{it} = 1 \\ v\left(e_{it}, x_{i,t+1}, 1\right) + \gamma - \ln P\left(d=1 | e_{it}, x_{i,t+1}\right) & \text{if } d_{it} = 0 \end{cases}$$

- Eliminate continuation values after intervention

$$\begin{aligned} v\left(e_{it}+1, x_{i,t+1}, 0\right) &= \theta \ln\left(y_{i,t+1}\right) + \beta E_x\left[\overline{V}\left(e_{it}+1, x_{i,t+2} : \theta\right) | x_{i,t+1}\right] \\ v\left(e_{it}, x_{i,t+1}, 1\right) &= \theta \ln\left(y_{i,t+1} - s_{i,t+1}\right) + \beta E_x\left[\overline{V}\left(e_{it}+1, x_{i,t+2} : \theta\right) | x_{i,t+1}\right]. \end{aligned}$$

- Substitute back to HM inversion

Back

CCP estimation

- Estimate probability of schooling at each state value
  - Needed to construct dependent variable in 2SLS regression
- Smooth probabilities across states by using a flexible logit
  - Ideally frequency estimates for each state
  - Practically no variation for some states
- Choose MLE or GMM to replicate treatment effects at this stage

Back

Robustness to how to extrapolate across contexts

		Across-contexts	
	Target	HTE	PSW
Treatment	0.0567***	0.212***	0.00660
	(0.0106)	(0.00442)	(0.0184)
Control mean	0.894***	1.127***	0.895***
	(0.00951)	(0.00373)	(0.0128)
Obs.	4982	4982	1490
= Target TE		0.000	0.007
= Target control mean		0.000	0.927

Note: *** p<0.01 ** p<0.05 * p<0.1

Back

Robustness to how to extrapolate $\Delta R(e;x)$

		Across-contexts		Across-policies	
	Target	Linear	RF	Linear	RF
Treatment	0.0567***	0.0431***	0.0412***	0.0590***	0.0577***
	(0.0106)	(0.00465)	(0.00644)	(0.00545)	(0.00542)
Control mean	0.894***	0.702***	0.676***	0.941***	0.942***
	(0.00951)	(0.00390)	(0.00556)	(0.00531)	(0.00529)
Obs.	4982	4982	4982	4982	4982
= Target TE		0.004	0.016	0.674	0.863
= Target control mean		0.000	0.000	0.000	0.000

Note: *** p<0.01 ** p<0.05 * p<0.1

Back

Comparison of extrapolation methods

	Target	Across-contexts			
		HTE		PSW	
Treatment	0.0567***	0.212***	0.0308***	0.00660	-0.0459**
	(0.0106)	(0.00442)	(0.00421)	(0.0184)	(0.0202)
Control mean	0.894***	1.127***	1.111***	0.895***	1.000***
	(0.00951)	(0.00373)	(0.00367)	(0.0128)	(0.00221)
Obs.	4982	4982	4982	1490	1490
= Target TE		0.000	0.000	0.007	0.000
= Target control mean		0.000	0.000	0.927	0.000
Normalization of s, y, z			✓		✓

Note: *** p<0.01 ** p<0.05 * p<0.1

- Structural: school costs & cash transfers **relative to per-capita income**
- Across-contexts using relative values in reduced-form not improved

Back

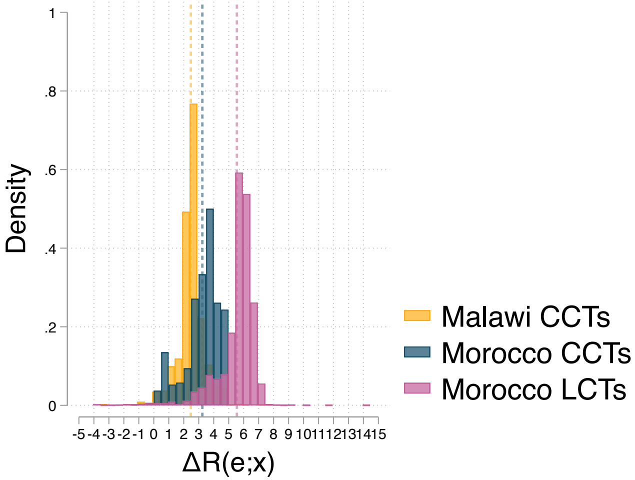
Estimates of elasticity of schooling

	Malawi	Morocco	
	CCTs	CCTs	LCTs
$E\left[\frac{\partial P^1_{i,2}}{\partial z_{i,2}} \frac{z_{i,2}}{P^1_{i,2}}\right]$	0.185	0.332	0.142

- $E\left[\frac{\partial P^1_{i,2}}{\partial z_{i,2}} \frac{z_{i,2}}{P^1_{i,2}}\right]$: average elasticity of schooling w.r.t cash transfers
- Across-contexts: difference in effective size of cash transfer
- Across-policies: substitution effects

Back

Estimates of $\Delta R_i(e;x)$



- Dashed lines indicate $E[\Delta R(e;x)]$ under each experiment

Model fit

	Malawi	Morocco	
	CCTs	CCTs	LCTs
Treatment	0.0317***	0.0554***	0.0539***
	(0.00495)	(0.00777)	(0.00954)
Control mean	0.895***	0.894***	0.900***
	(0.00241)	(0.00747)	(0.00889)
Obs.	1490	4982	3018
= Target TE	0.290	0.869	0.051
= Target control mean	0.721	0.981	0.476

Note: *** p<0.01 ** p<0.05 * p<0.1

Back

Extrapolation from Malawi CCTs with true θ or $\Delta R(e;x)$

	Target	Across-contexts		
Treatment	0.0567***	0.0431***	0.0953***	0.0367***
	(0.0106)	(0.00465)	(0.00486)	(0.00721)
Control mean	0.894***	0.702***	0.688***	0.901***
	(0.00951)	(0.00390)	(0.00411)	(0.00689)
Obs.	4982	4982	4982	4982
Replace θ			✓	
Replace $\Delta R(e;x)$				✓
= Target TE		0.004	0.000	0.006
= Target control mean		0.000	0.000	0.276

Note: *** p<0.01 ** p<0.05 * p<0.1

Back

Within-intervention variation of $\Delta R(e;x)$

	$E[\theta \ln(y - s/y)]$		$E[\beta \Delta R(e;x)]$	
	Control	Treatment	Control	Treatment
Malawi CCTs	-0.292	0.594***	2.526	2.373*
Morocco CCTs	-0.117	0.409***	3.045	3.085
Morocco LCTs	-1.706	-1.390***	5.109	5.383***

*** p<0.01 ** p<0.05 * p<0.1 for difference across groups in each intervention.

- Within-intervention variation affects prediction of treatment effect
- Malawi CCTs: $\Delta R(e;x)$ smaller for treatment group
  - Children with lower perceived returns to choose schooling
- Moroccan CCTs: $\Delta R(e;x)$ larger for treatment group (if anything)
  - Same pattern for Moroccan LCTs

Back

Reduced-form extrapolation with normalization

	Target	Across-contexts		
		Structural	HTE	PSW
Treatment	0.0567***	0.0150***	0.0352***	0.0655***
	(0.0106)	(0.000788)	(0.00346)	(0.0140)
Control mean	0.894***	0.903***	0.894***	0.920***
	(0.00951)	(0.000674)	(0.00302)	(0.0112)
Obs.	4982	4982	4982	1490
= Target TE		0.000	0.000	0.534
= Target control mean		0.000	0.833	0.022

Note: *** p<0.01 ** p<0.05 * p<0.1

- HTE & PSW: $s/y, z/y$ & standardized e and age

Back

Confusion about conditionality of Moroccan CCTs

- Moroccan CCTs sample largely misunderstood conditionality
  - 11% understood conditionality correctly
  - 14% thought CCTs were tied to school enrollment
  - More similar to Moroccan LCTs than Malawi CCTs

- If perfect understanding, across-contexts extrapolation more accurate?

- Compute counterfactual effect of Moroccan CCTs with no confusion
  - Estimate model under Moroccan CCTs with degree of confusion
  - Simulate model by assuming perfect knowledge

	Estimation		Across-contexts	Across-policies
	Original	Counterfactual	Linear	Linear
Treatment	0.0567***	0.122***	0.0431***	0.0590***
	(0.0106)	(0.00836)	(0.00465)	(0.00545)
Control mean	0.894***	0.868***	0.702***	0.941***
	(0.00951)	(0.00814)	(0.00390)	(0.00531)
Obs.	4982	4982	4982	4982
= Target TE			0.000	0.000
= Target control mean			0.000	0.000

Note: *** $p<0.01$ ** $p<0.05$ * $p<0.1$

- Treatment effect becomes bigger under perfect knowledge
- Both extrapolations are statistically different from estimate

Back