How Transparency and Reproducibility Can Increase Credibility in Policy Analysis

A Case Study of the Minimum Wage Policy Estimates

Fernando Hoces de la Guardia

IADB, March 3, 2017

Motivation: Major Push In Science to Increase Quality Through Transparency and Reproducibility

Credible research is key step to achieve the ideal of common evidence across different policy makers.

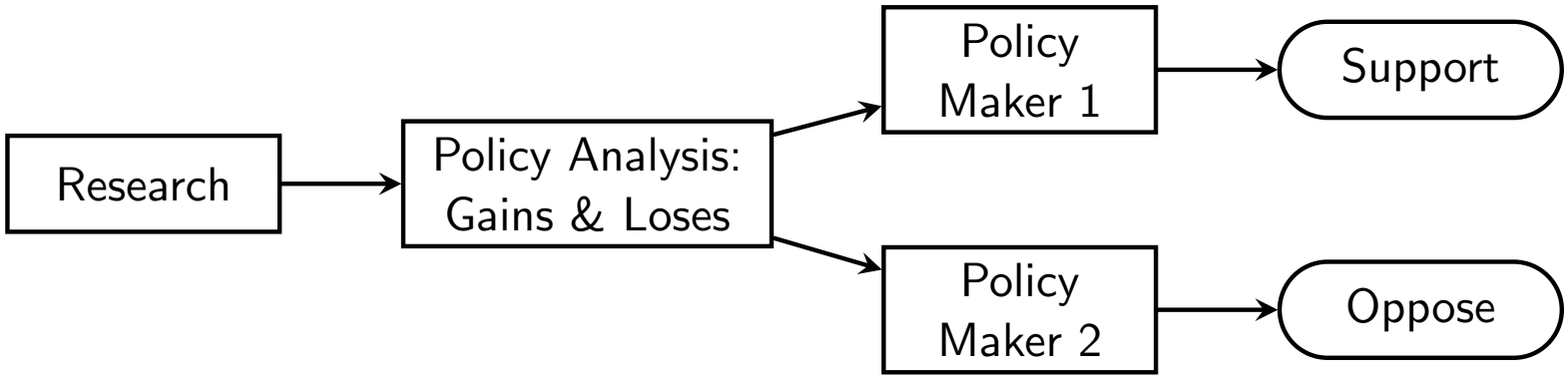


Figure: Simplified Model of Connection Between Evidence and Policy

Next step: increase the credibility of policy analysis.

Problem: Policy Analysis Often Has Low Credibility [Manski, 2013]

- **The Problem:** Little is known about the boundaries of the point estimates provided in policy reports.
- **Manski Proposes:** In increasing order of desirability: (1) display standard errors; (2) bound estimated effects; (3) add [policy] decision criteria to the analysis (best).

Problem: Policy Analysis Often Has Low Credibility [Manski, 2013]

- **The Problem:** Little is known about the boundaries of the point estimates provided in policy reports.
- **Manski Proposes:** In increasing order of desirability: (1) display standard errors; (2) bound estimated effects; (3) add [policy] decision criteria to the analysis (best).

Contribution: Bring Open Science To Policy Analysis

As a complement to Manski's prescription, this paper incorporates guidelines and tools created in response to the reproducibility crisis in science into policy analysis. This methodological development is implemented in the case study.

Lack of transparency and reproducibility in policy analysis

Why does it matter?

- Threatens the credibility of the policy estimates
- Difficult to understand precisely how research affects policy estimates
- Hinders automation and/or systematic improvements of reports

The Transparent and Reproducible (TR) Approach to Policy Analysis

1. Comparison of Credibility Crises in Research and Policy Analysis
2. Translation of Guidelines and Tools to The Policy Analysis Setting
3. Application to Case Study: Policy Analysis on Minimum Wage
   - Demonstrate how to achieve highest standards of TR
   - Use sensitivity analysis to explore biggest policy unknowns
   - Surprisingly academic debate around one specific parameter seems less relevant from policy perspective

Comparison of Low Transparency and Reproducibility (TR) in Research and Policy Analysis

	Research	Policy Analysis
Output	Peer review publication	Policy report
Problems of low TR	Publication Bias. Specification Search (P-Hacking, Garden of forking paths). Data fudging.	Low credibility. Unclear connection between research and policy. Hard to improve systematically. Data fudging.
Common Solutions	Disclosure of key details. Open data and materials.	
Common Tools	Dynamic documentation. Distributed version control.	
Specific Solutions	Test for reproducibility; Registration of pre-analysis	Develop reproducibility; Systematic and continuous updating
Who increases TR	Researchers, Funders, Journals	**Not** the policy analysts (Policy schools? Think tanks? Discussion at the end)

We don't know how the sausage is made!

We don't know how the sausage is made!

Let's follow science, open up the kitchen.
And publish the cook book with the recipe.

Description of Case study

"The Effects of a Minimum-Wage Increase on Employment and Family Income" Congressional Budget Office (2014)

Description: CBO estimated the effects of a raise in the federal minimum wage from \$7.25/hr to \$10.10/hr.

Main policy estimates:

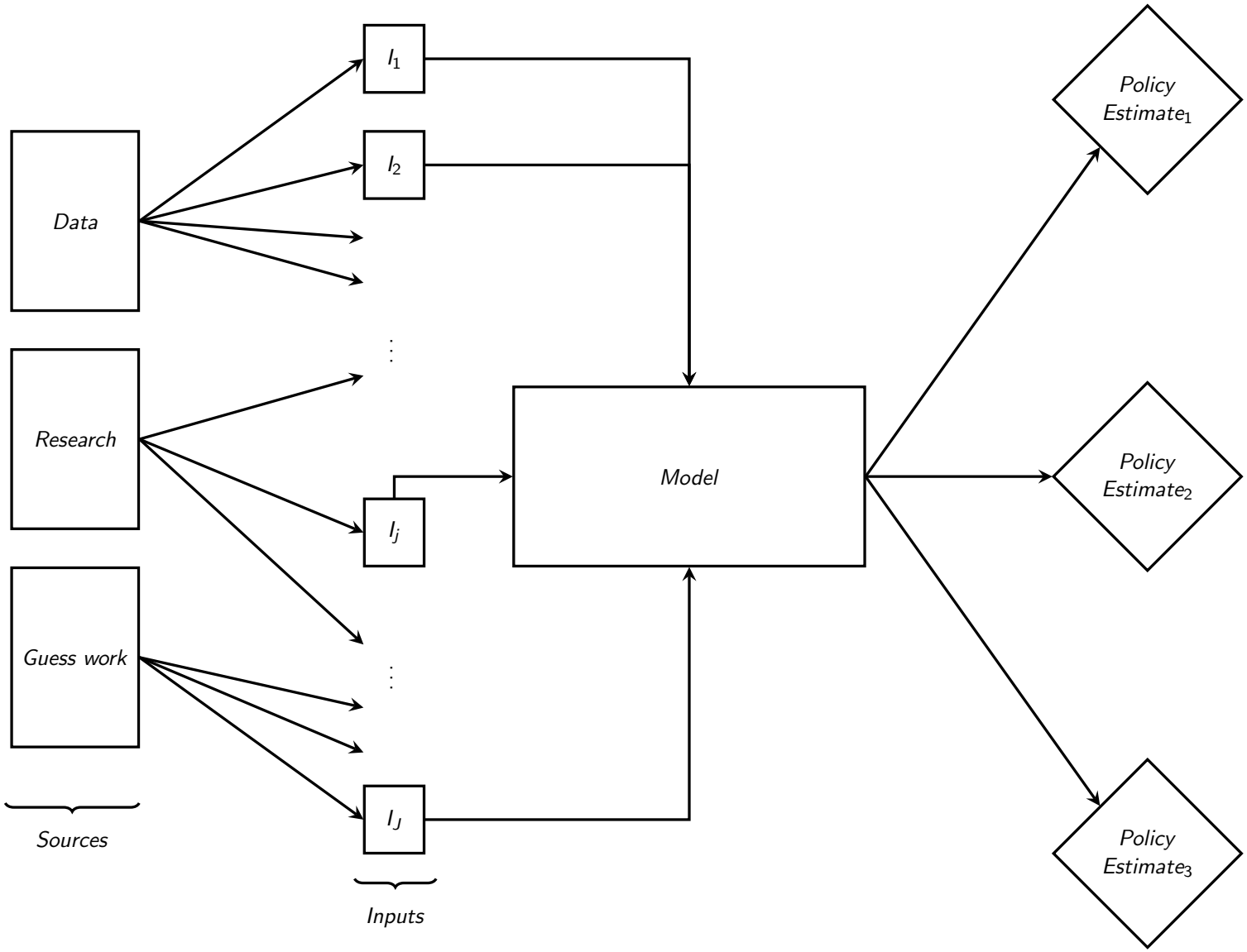
- 500,000 jobs would be lost.
- 16.5 million workers would receive a salary increase.
- Distributional effects: below poverty line (PL) +\$5billion; between one and three PL +\$12billion; between three and six PL +\$2billion; above six PL -\$17billion

Key research estimate: Elasticity of labor demand.

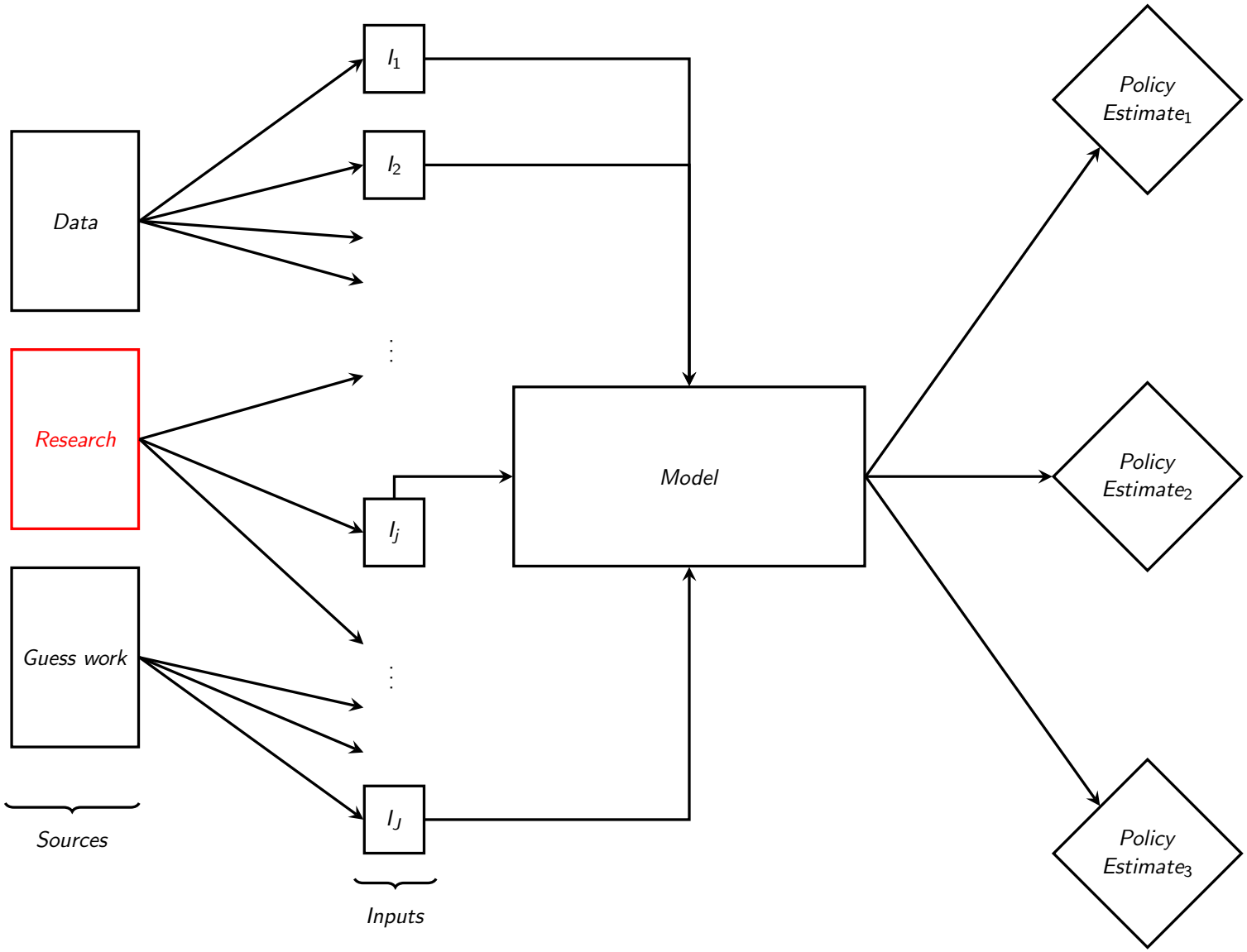
Results

1. Guidelines for TR Policy Analysis
2. Application to Case Study: Policy Analysis on Minimum Wage
   - Score the current report
   - Use guidelines and tools demonstrate high TR
   - Perform sensitivity analysis to identify biggest gaps in (policy relevant) knowledge

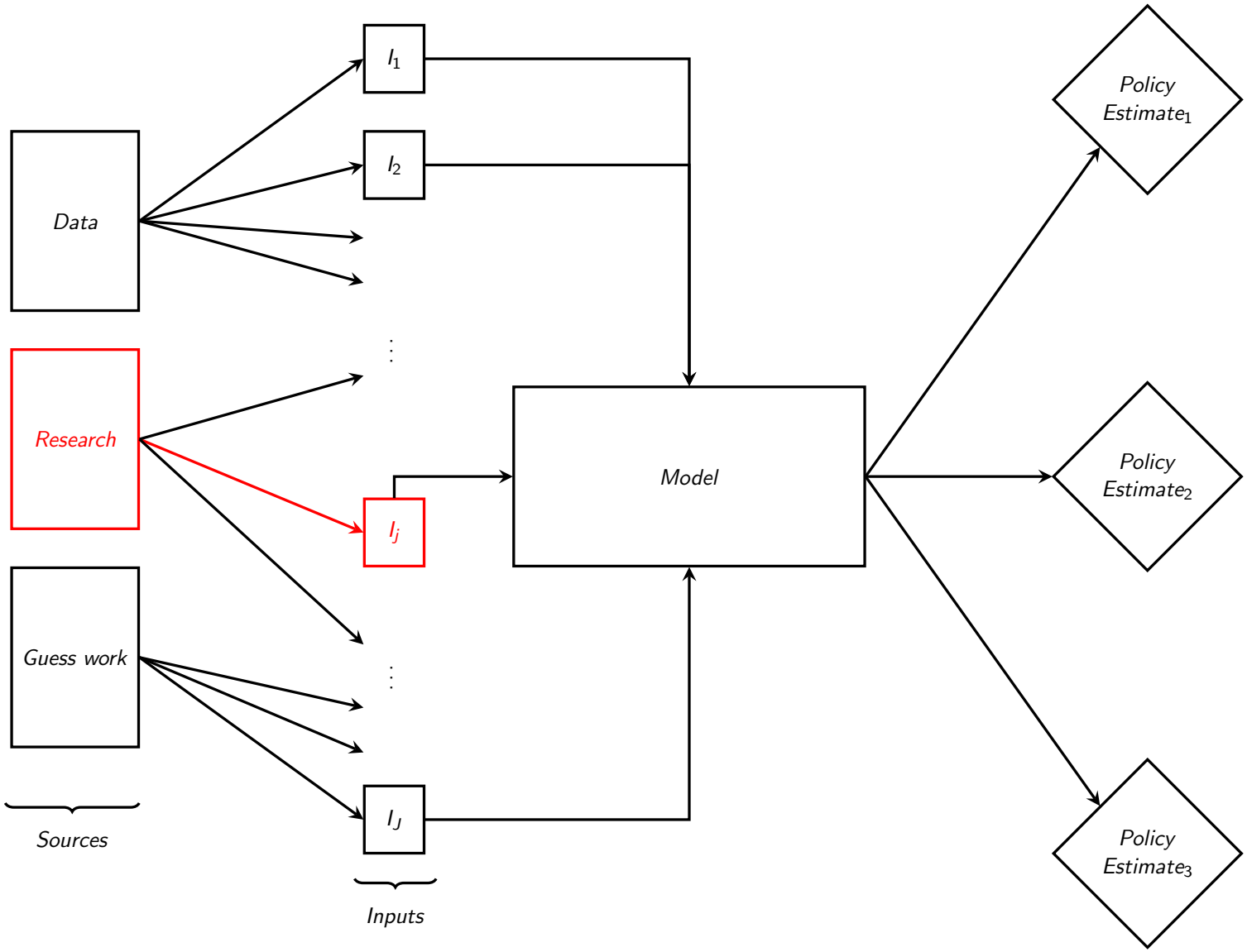
Guidelines Goal: Map the complete policy analysis. Transparent & Reproducible



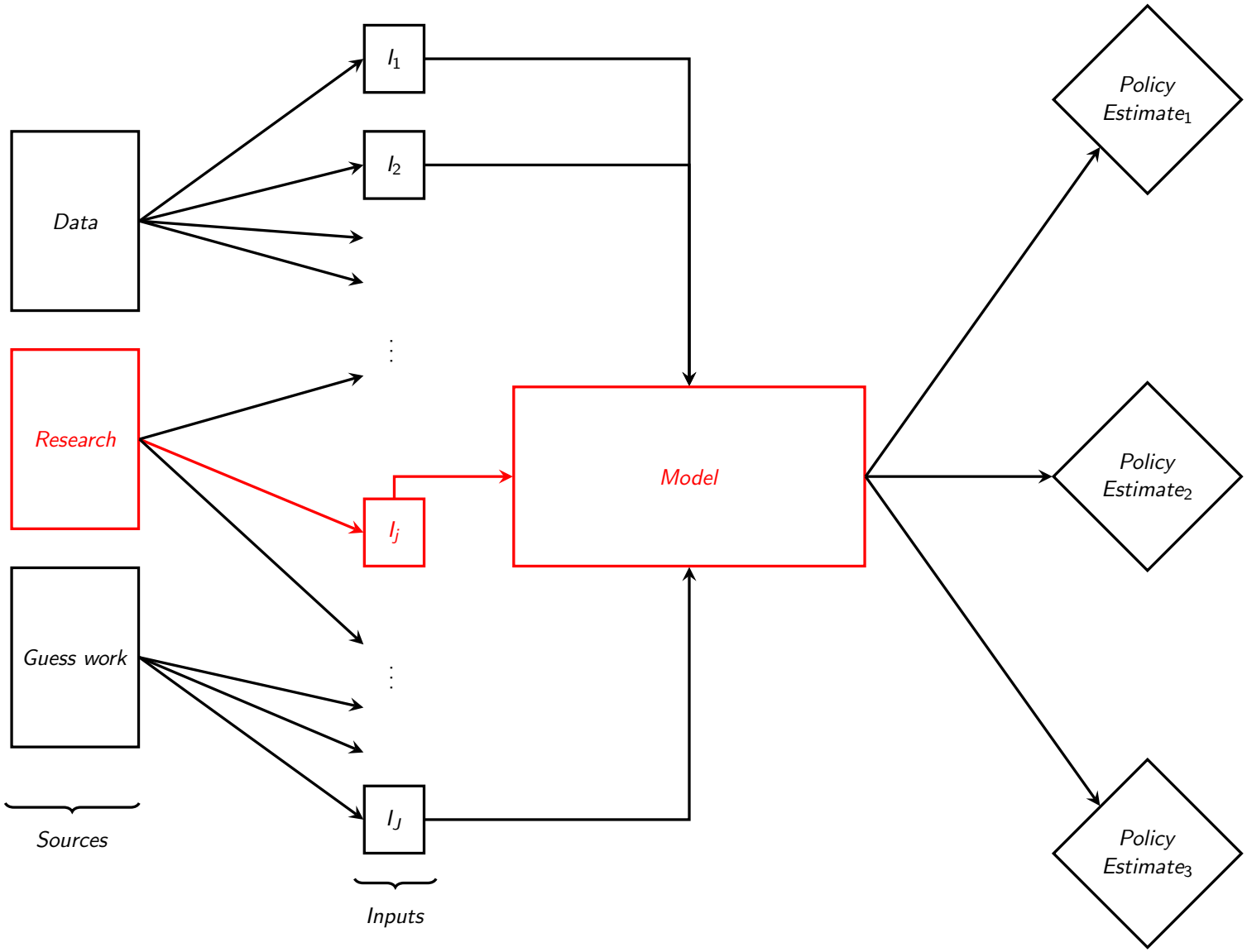
Guidelines Goal: Map the complete policy analysis. Transparent & Reproducible



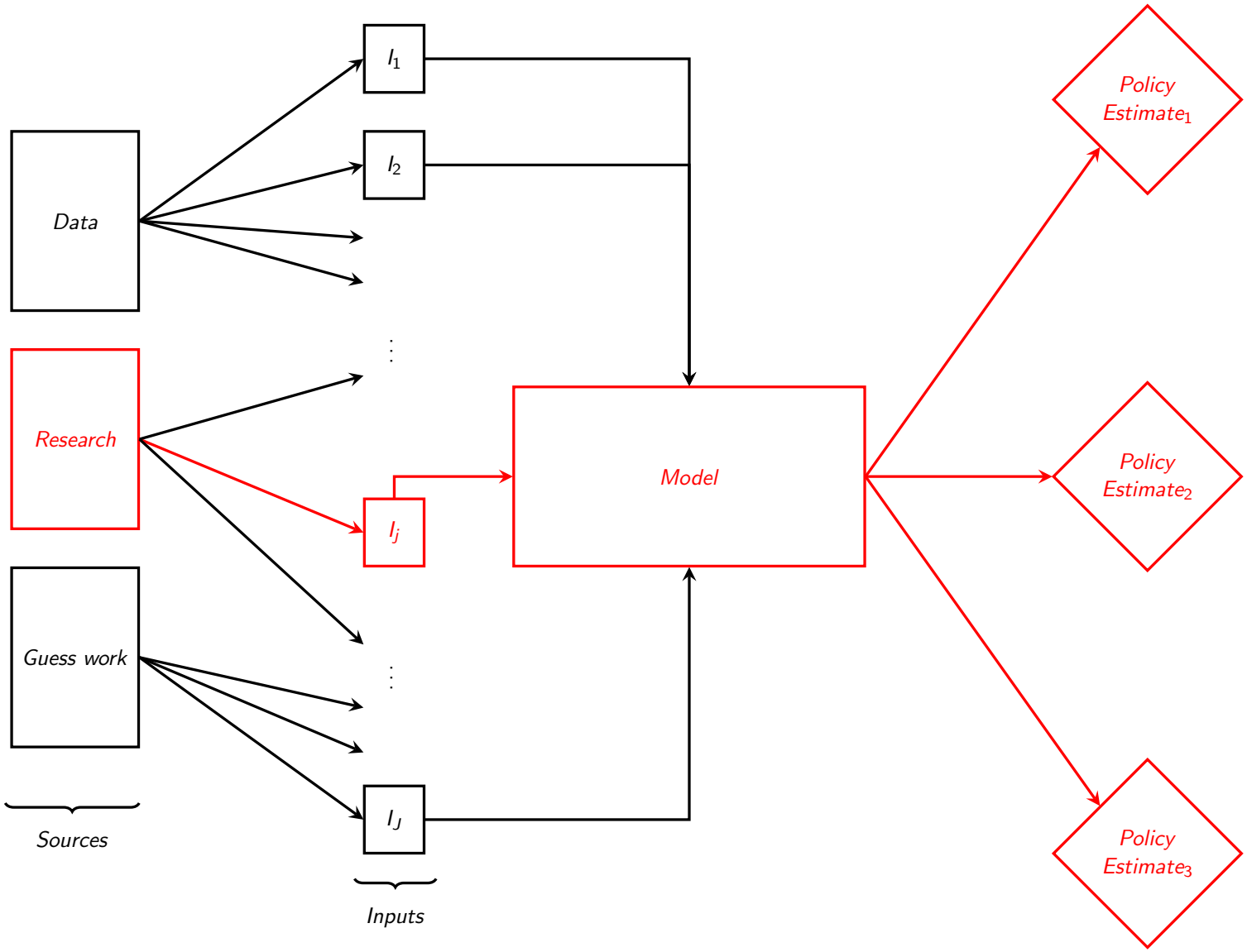
Guidelines Goal: Map the complete policy analysis. Transparent & Reproducible



Guidelines Goal: Map the complete policy analysis. Transparent & Reproducible



Guidelines Goal: Map the complete policy analysis. Transparent & Reproducible



Summary of Adapted Guidelines

Standard	Level 0	Level 1	Level 2	Level 3
Sources (Data, Research, Guesswork)	Report says nothing	Clearly stated whether all, some components, or none of the data is available, with instructions for access when possible.	Lvl 1 + report and data are in same place	Lvl 2 + Code embedded in the report calls the data and changes in the data produce traceable changes in the report
Methods & Code	Key assumption are listed	Methods are described in prose. Large amount of work is required to reproduce qualitatively similar estimates	Methods and described in prose, with detailed formulas, and code is provided as supplementary material	Lvl 2 + All is in the same document where changes in the code affect the output automatically
Workflow	Policy estimates vaguely described	All the inputs, and their corresponding sources, used in the calculations are listed	Lvl 1 + Policy estimates are listed, in same unit if possible	Lvl 2 + all the components can be modified with little effort

From TOP guidelines [Nosek et al., 2015]

Applying Guidelines to CBO Report

Standard	Level 0	Level 1	Level 2	Level 3
Sources (Data, Re-search, Guess-work)	Report says nothing	Clearly stated whether all, some components, or none of the data is available, with instructions for access when possible.	Lvl 1 + report and data are in same place	Lvl 2 + Code embedded in the report calls the data and changes in the data produce traceable changes in the report
Methods & Code	Key as-sumption are listed	Methods are described in prose. Large amount of work is required to reproduce qualitatively similar estimates	Methods and described in prose, with detailed formulas, and code is provided as supplementary material	Lvl 2 + All is in the same document where changes in the code affect the output automatically
Work-flow	Policy estimates vaguely described	All the inputs, and their corresponding sources, used in the calculations are listed [80%]	Lvl 1 + Policy estimates are listed, in same unit if possible	Lvl 2 + all the components can be modified with little effort

Adapted from TOP guidelines

Example: Verifying dis-employment by looking for elasticity values in report. **Before**

Steps taken to verify the analysis & employment variation ($\widehat{\Delta E} \times 1000$) at each line[1]

[1] Assuming target population $\approx$ 22 million, $\overline{\Delta w_{w \leq MW'}} \approx 14\%$, and non-compliance $\approx 15\%$

Example: Verifying dis-employment by looking for elasticity values in report. **Before**

Steps taken to verify the analysis & employment variation ($\widehat{\Delta E} \times 1000$) at each line[1]

1. Find an elasticity: -0.1 (page 25): $\widehat{\Delta E} \approx 300$

[1] Assuming target population $\approx$ 22 million, $\overline{\Delta w_{w \leq MW'}} \approx 14\%$, and non-compliance $\approx 15\%$

Example: Verifying dis-employment by looking for elasticity values in report. **Before**

Steps taken to verify the analysis & employment variation ($\widehat{\Delta E} \times 1000$) at each line[1]

1. Find an elasticity: -0.1 (page 25): $\widehat{\Delta E} \approx 300$
2. What about adults? $\eta^{adults} = \frac{1}{3}\eta^{teens}$ (page 28): $\widehat{\Delta E} \approx 100$

[1] Assuming target population $\approx$ 22 million, $\overline{\Delta w_{w \leq MW'}} \approx 14\%$, and non-compliance $\approx 15\%$

Example: Verifying dis-employment by looking for elasticity values in report. **Before**

Steps taken to verify the analysis & employment variation ($\widehat{\Delta E} \times 1000$) at each line[1]

1. Find an elasticity: -0.1 (page 25): $\widehat{\Delta E} \approx 300$
2. What about adults? $\eta^{adults} = \frac{1}{3}\eta^{teens}$ (page 28): $\widehat{\Delta E} \approx 100$
3. What about the adjustment? $\widetilde{\eta^{g}_{w \leq MW}} = \frac{\eta^{g}_{lit}}{p^{g}_{w \leq MW}} \times \frac{\%\Delta MW}{\%\Delta w^{g}}$ (page 26-28 + 2 papers): $\widehat{\Delta E} \approx 1,100$

[1] Assuming target population $\approx$ 22 million, $\overline{\Delta w_{w \leq MW'}} \approx 14\%$, and non-compliance $\approx 15\%$

Example: Verifying dis-employment by looking for elasticity values in report. **Before**

Steps taken to verify the analysis & employment variation ($\widehat{\Delta E} \times 1000$) at each line[1]

1. Find an elasticity: -0.1 (page 25): $\widehat{\Delta E} \approx 300$
2. What about adults? $\eta^{adults} = \frac{1}{3}\eta^{teens}$ (page 28): $\widehat{\Delta E} \approx 100$
3. What about the adjustment? $\widetilde{\eta^{g}_{w \leq MW}} = \frac{\eta^{g}_{lit}}{p^{g}_{w \leq MW}} \times \frac{\%\Delta MW}{\%\Delta w^{g}}$ (page 26-28 + 2 papers): $\widehat{\Delta E} \approx 1,100$
4. The adjustment factors $\frac{1}{p^{g}_{w \leq MW}} \times \frac{\%\Delta MW}{\%\Delta w^{g}} = F^{g}_{adj}$ are not computed from the data (3.2 teens, 19.5 adults). Instead: $F^{teen}_{adj} = F^{adult}_{adj} = 4.5$ (page 28) $\widehat{\Delta E} \approx 500$

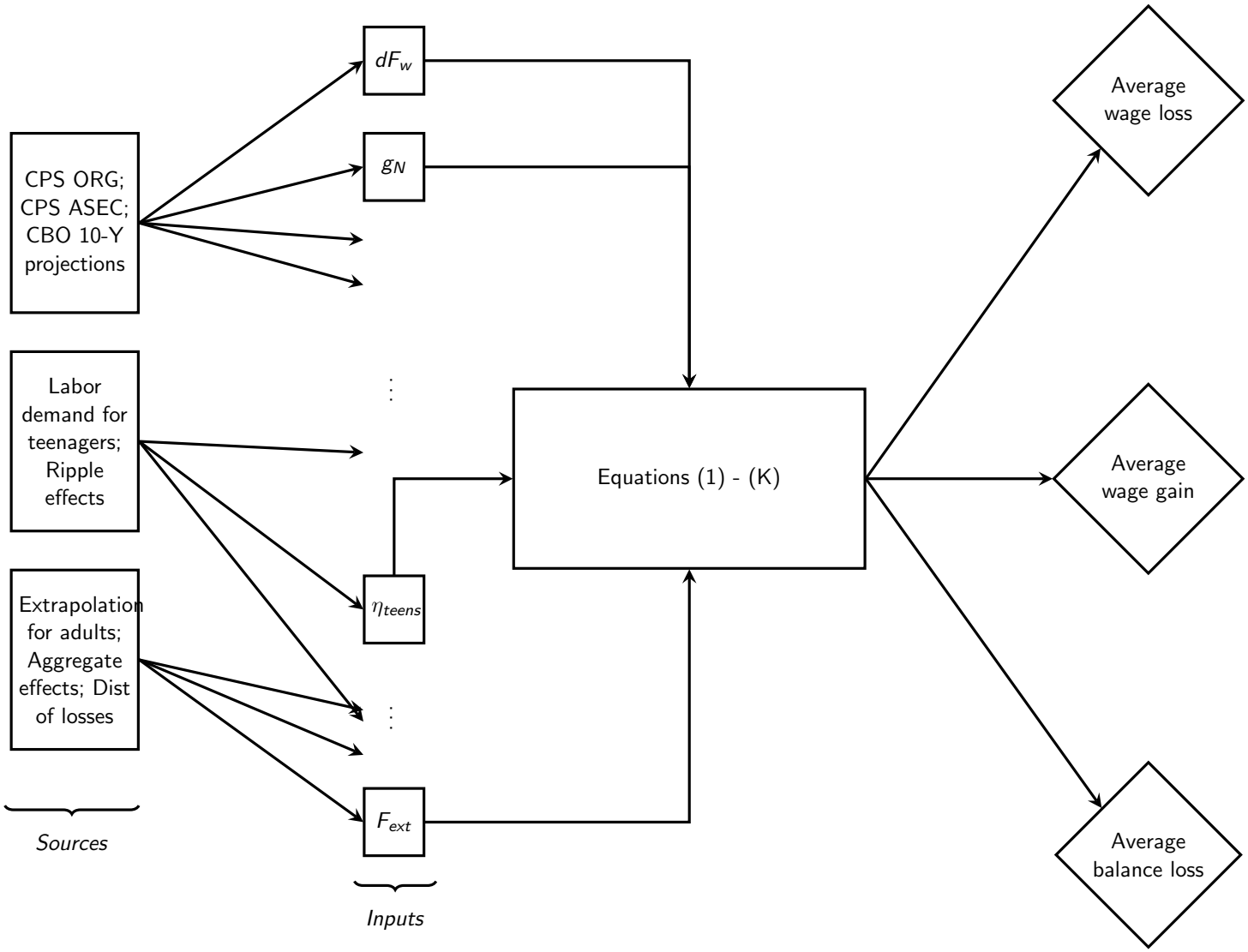
Steps 2-4 took several days of work!

[1] Assuming target population $\approx$ 22 million, $\overline{\Delta w_{w \leq MW'}} \approx 14\%$, and non-compliance $\approx 15\%$

After: Dynamic Document

DEMO .

Map the complete policy analysis. Example with Case Study



Clear Display of Policy Estimates: Step 1/3

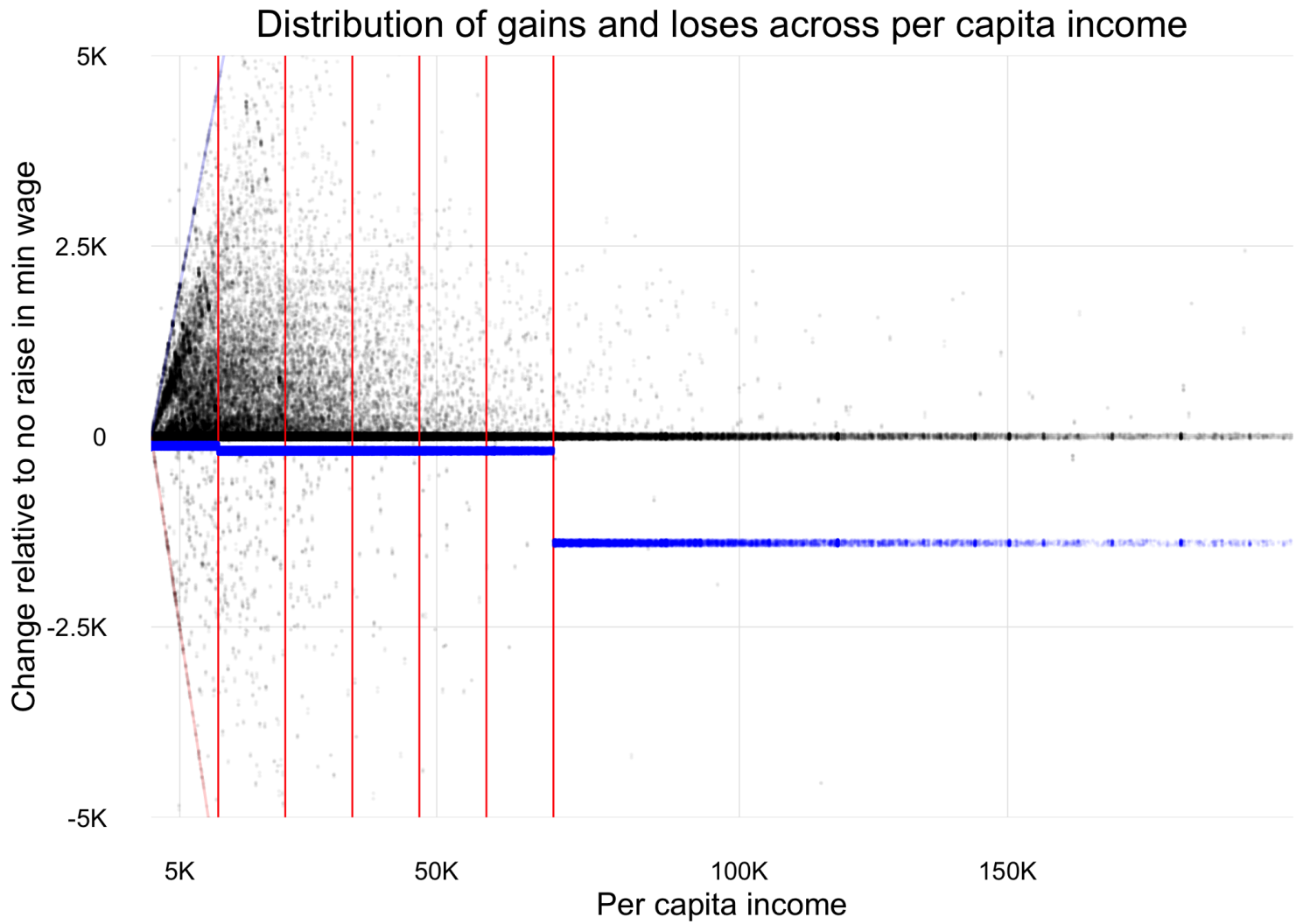


Figure: Gains and losses. Different Units

Clear Display of Policy Estimates: Step 2/3

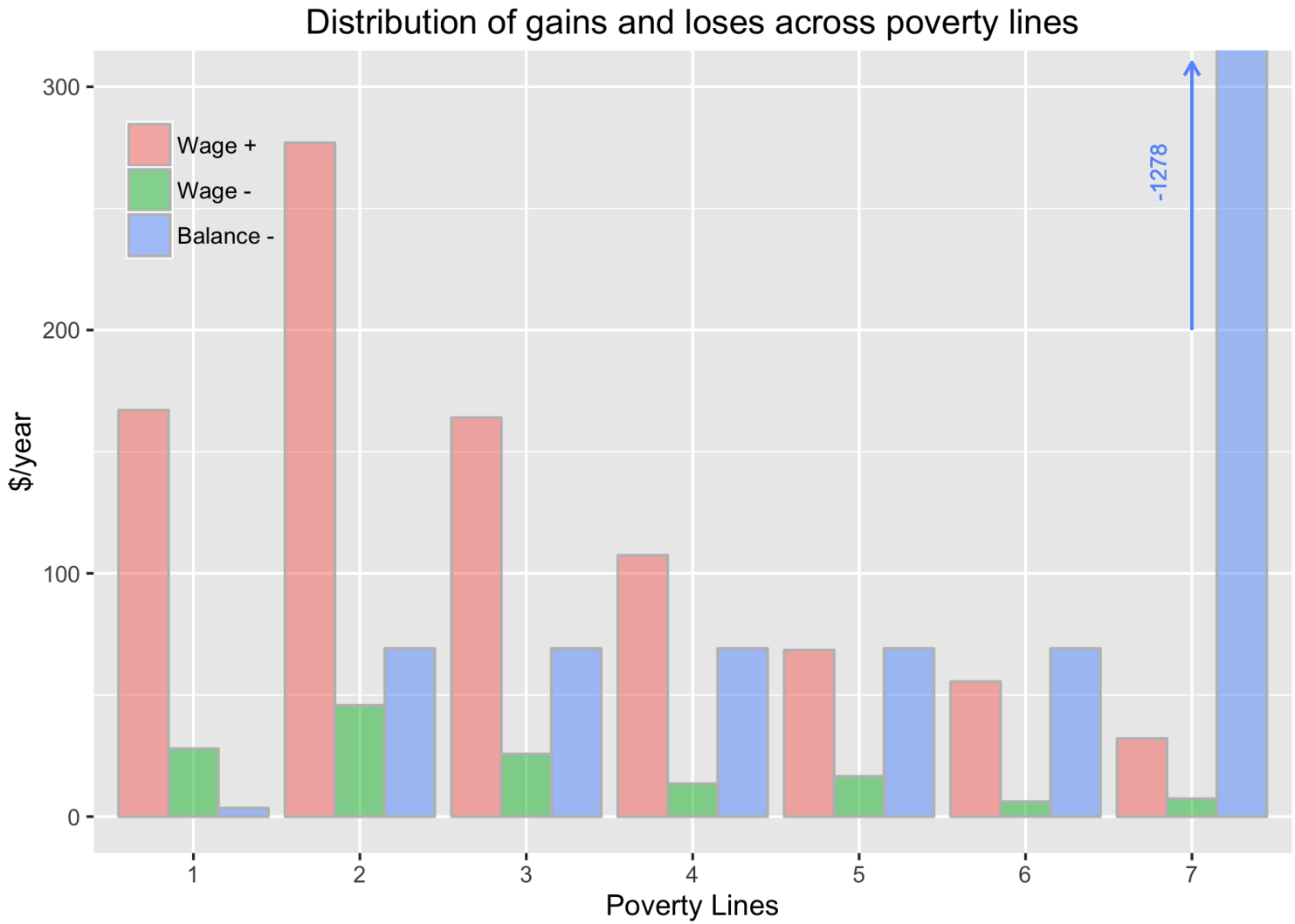


Figure: Gains and losses. Different Denominator

Clear Display of Policy Estimates: Step 3/3

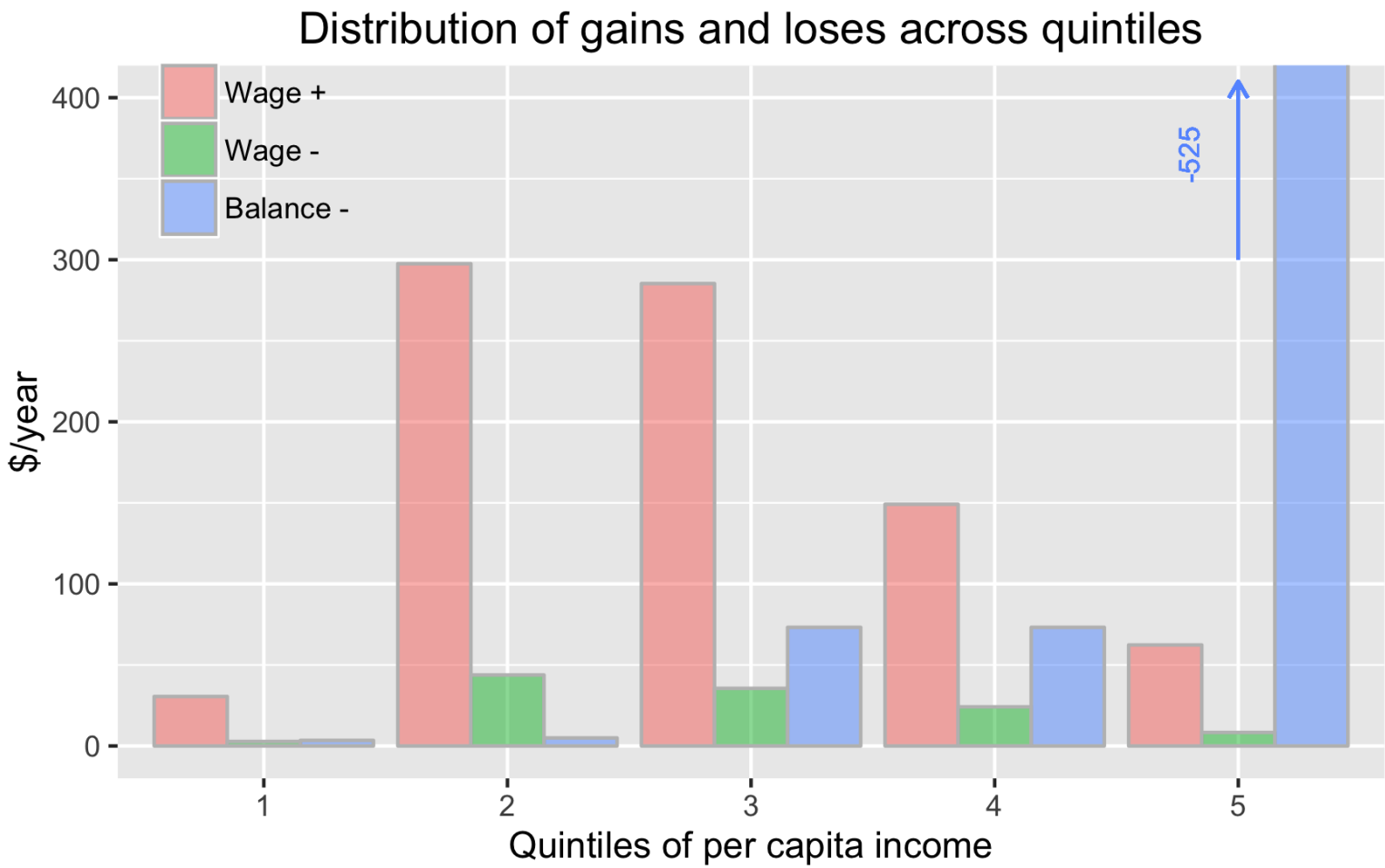


Figure: Gains and losses. Same units and denominator

Sensitivity Analysis: Status Quo

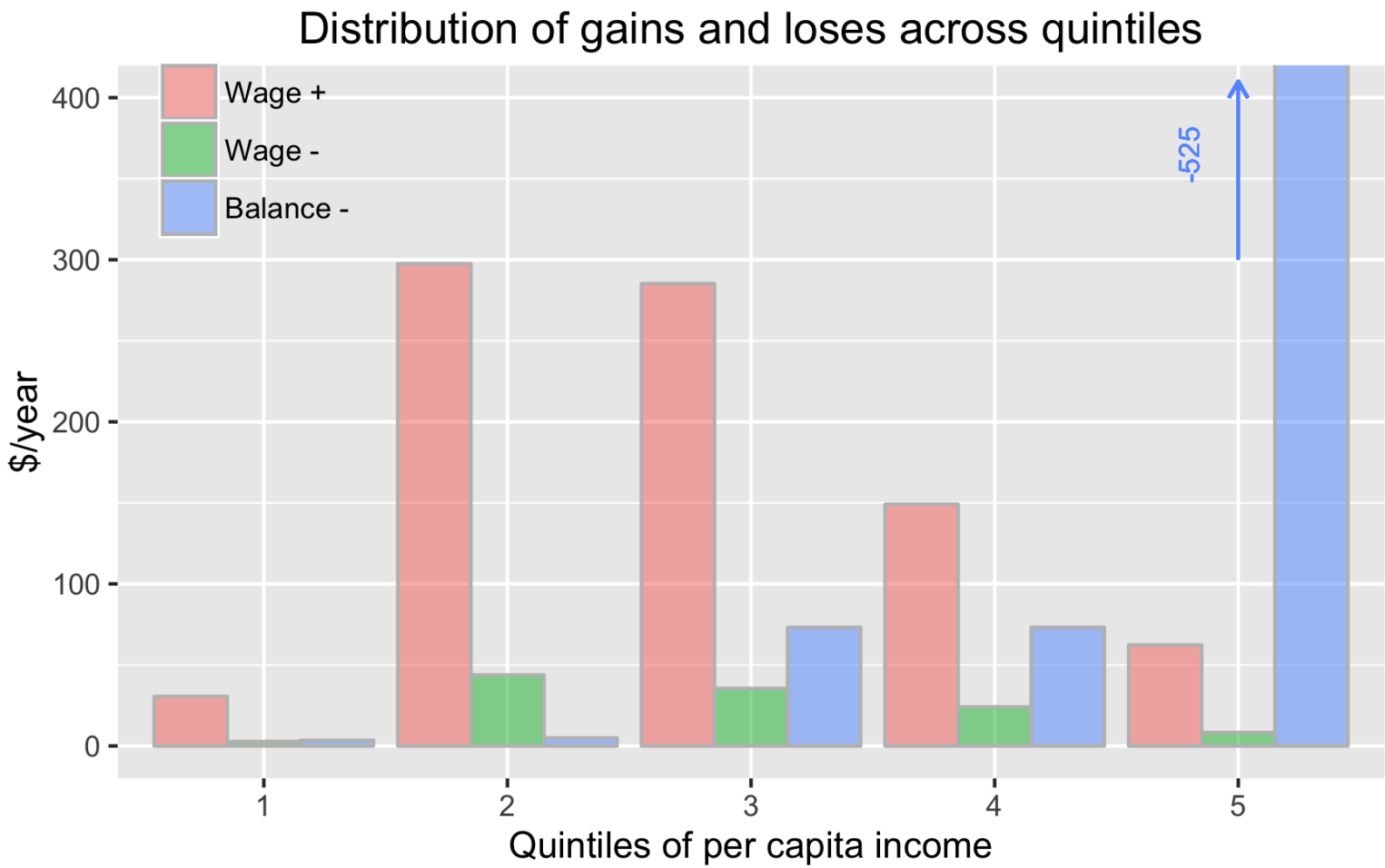


Figure: Default settings

SA: Change in Elasticity of Labor Demand

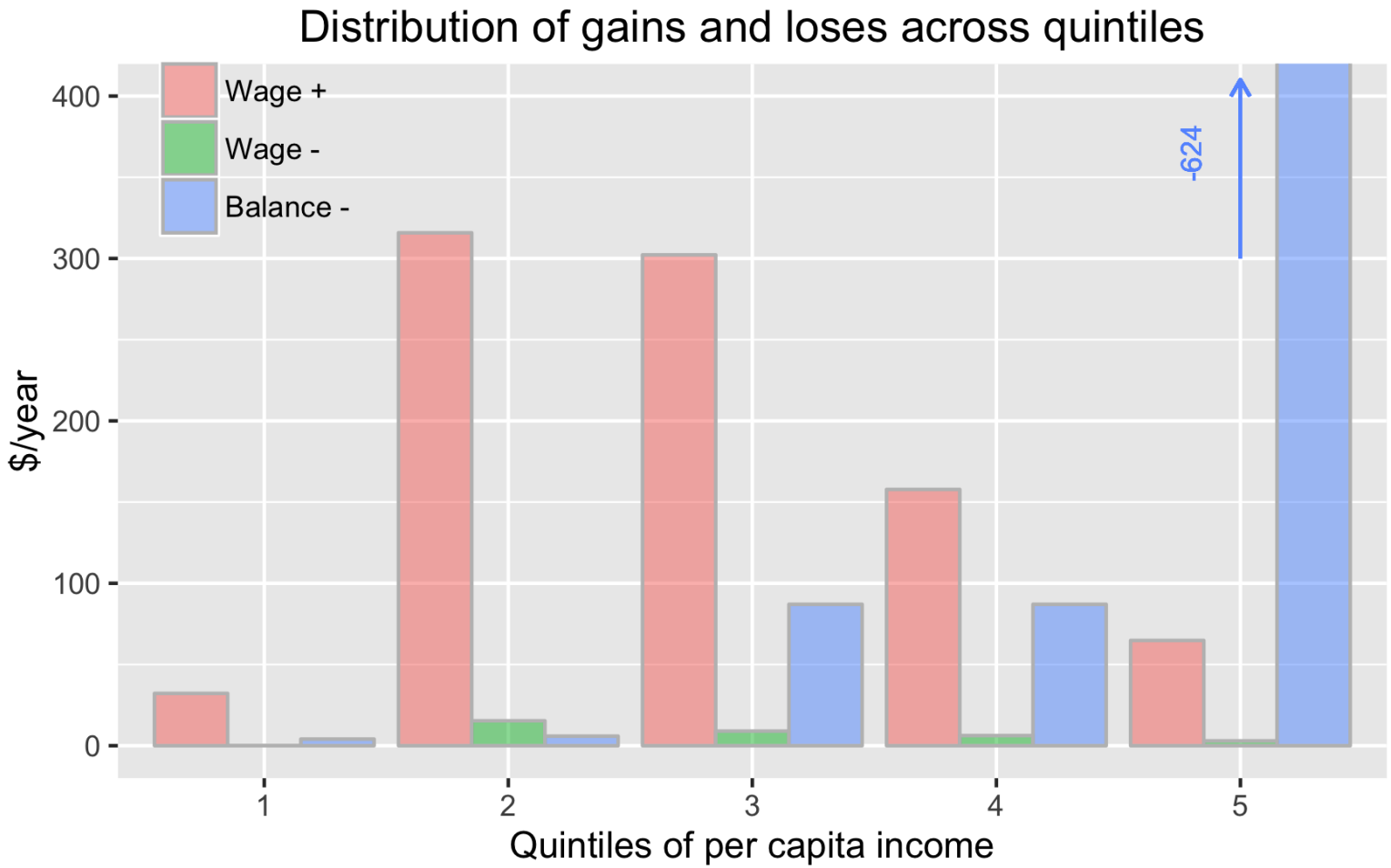


Figure: From $\eta_{lit}^{teens} = -0.1$ to $\eta_{lit}^{teens} = -0.01 (\Delta^{-} 90\%)$

Sensitivity Analysis: Status Quo

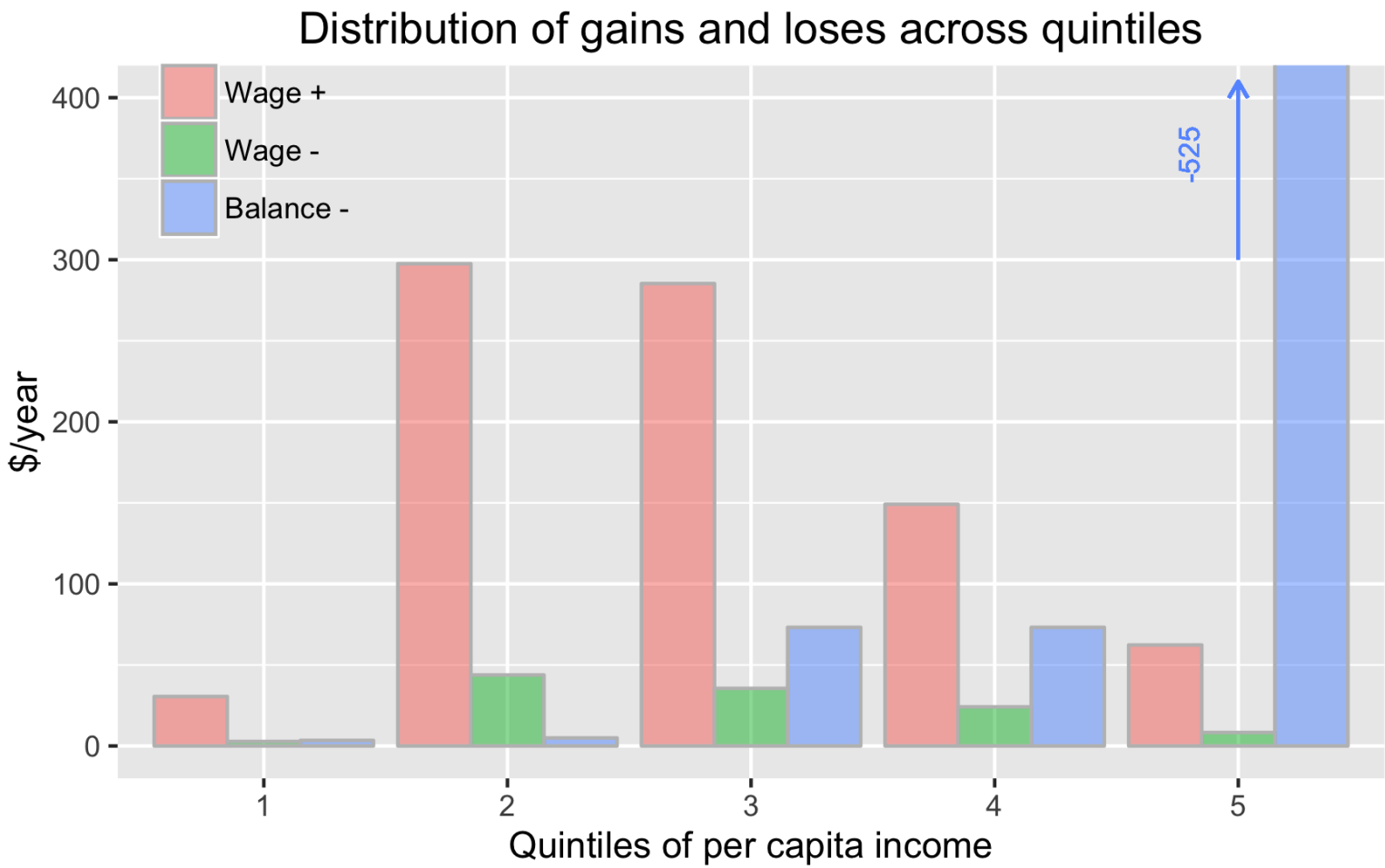


Figure: Default settings

SA: Change in Distribution of Balance Loses

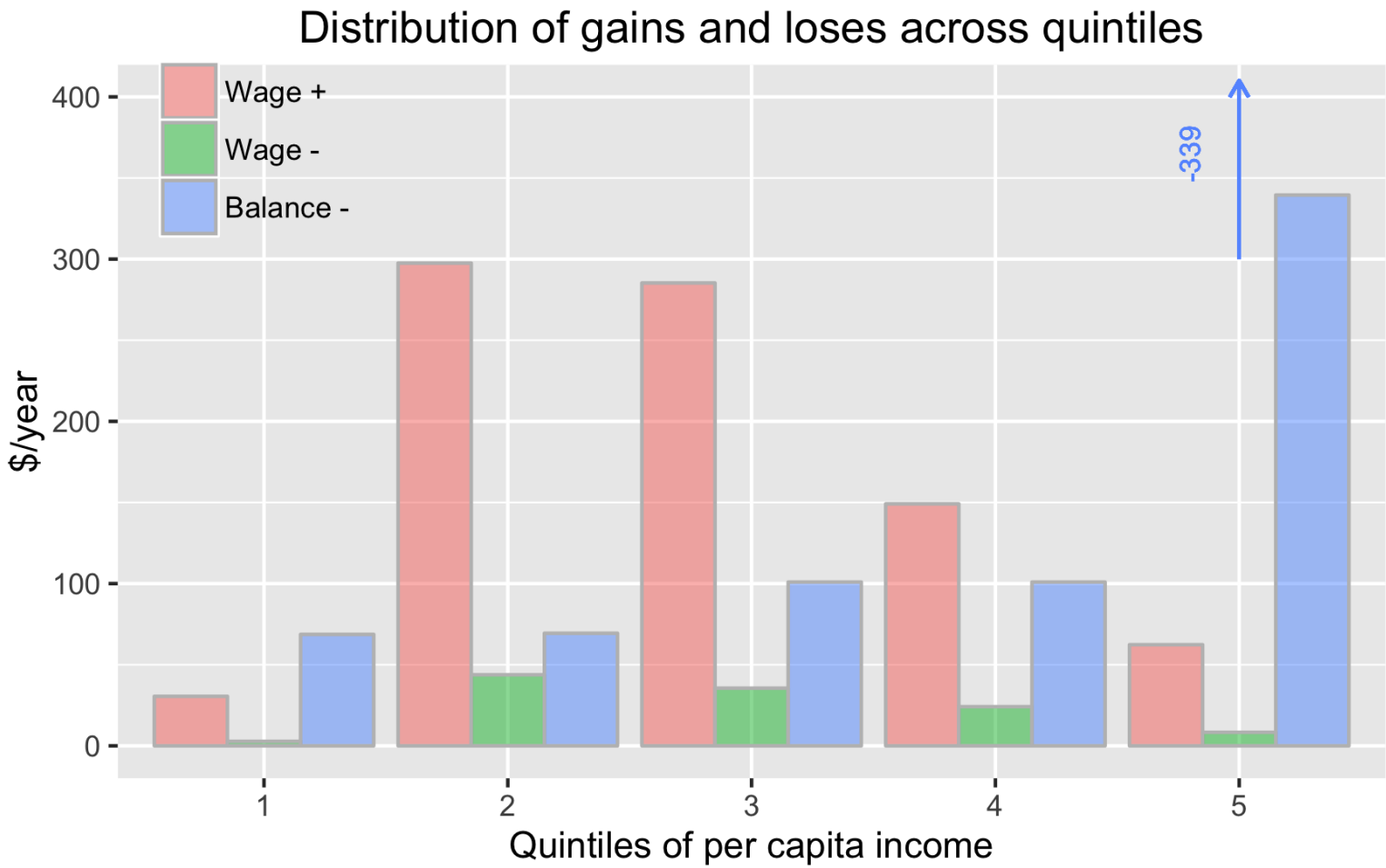


Figure: From $(1PL, 6PL) \sim (1\%, 29\%, 70\%)$ to $(20\%, 40\%, 40\%)$

Comparing the Trade-offs: A Toy Example

Model for the normative comparison made by a policy maker (welfare function):

$$W(\omega) = \sum_{i \in N} \left(\omega_{WG} WG_i + \omega_{WL} WL_i + \omega_{BL} BL_i\right) \omega_i^d(Q_i, \rho)$$

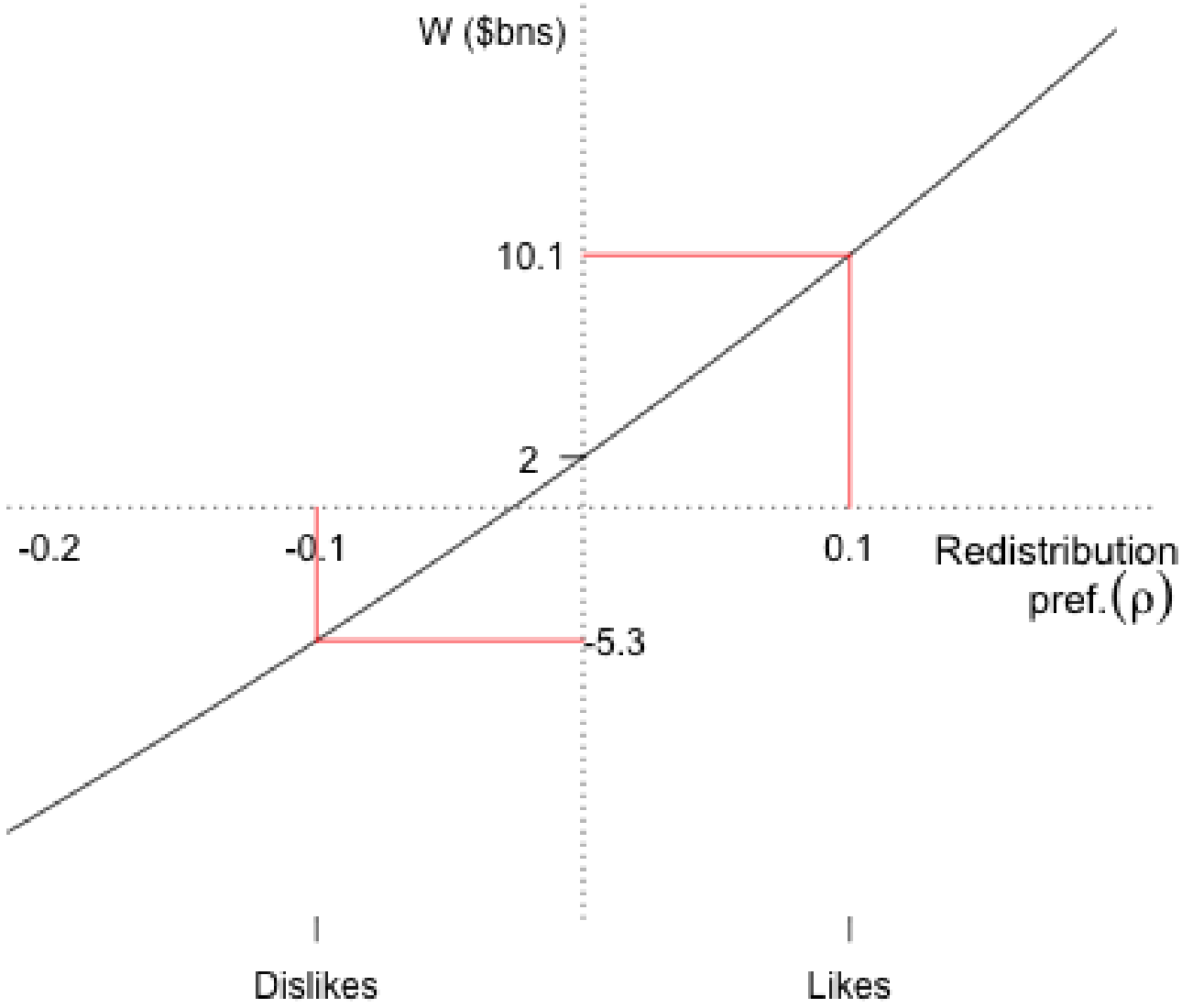
with:

$$\omega_i^d(Q_i, \rho) = \frac{(1 - \rho(Q_i - Q_{max}/2))}{\sum_i \omega_i^d(Q_i)} Q_{max} \quad \text{for } \rho > -\frac{2}{3}$$

$\rho > 0$ represent positive valuation of progressive redistribution.
$\rho < 0$ represents positive valuation of regressive redistribution (dis-utility from self loss greater than utility from others gain).

Normative Valuations and Redistribiutional Preferences

Toy Example ($\omega_{WG} = \omega_{WL} = \omega_{BL} = 1$)



Sensitivity Analysis For Multiple Parameters

Table: $\%\Delta W$ for a $\%\Delta$ in inputs. Two sample policy makers: dislikes ($W(-0.1) = -\$5.3bn$) and likes ($W(0.1) = \$10.1bn$) redistribution

		Redistributional Preferences			
		Dislikes ($\rho = -0.1$)		Likes ($\rho = 0.1$)	
Source	Input	$10\%\Delta^+$	$10\%\Delta^-$	$10\%\Delta^+$	$10\%\Delta^-$
Data					
	Annual wage growth (g_w)	-3%	2%	-2%	1%
	Annual growth in	0.8%	-0.9%	0.5%	-0.5%
	Wages in 2013 (w)	-38%	56%	-21%	32%
Research					
	η_{teen}	-4%	4%	-2%	2%
	Ripple Scope $(8.7, 11.5)$	37%	-24%	21%	-14%
	Ripple Intensity ($50\%\Delta w$) 5%	-5%	3%	-3%	
Guess Work					
	Extrapolation factor (F_{ex})	-3%	2%	-1%	1%
	Non compliance (α_1)	-7%	7%	-4%	4%
	Substitution factor (F_{sub})		20%		-8%
	Net benefits	-5%	5%	2%	-2%
	Distribution of balance loses				
	Current: $(1\%, 29\%, 70\%)$				
	$(1\%, 4\%, 95\%)$		22%		13%
	$(5\%, 35\%, 60\%)$		-17%		-9%
	$1/N$		-129%		-73%

Discussion

- Who should work on this:
  - Analytic reviewers of report; Research division within agencies; Study Commissions ("MWSC or MSWC?" [Card and Krueger, 2016])
  - Public policy schools.
  - Think tanks; Bank of knowledge [Clemens and Kremer, 2016].
- Two type of contributions (∼software development):
  - Short term: Within a given time period the model should be taken as given. Less freedom, but direct impact on the policy analysis.
  - Long term: Structural revisions occur in parallel and are incorporated in future cycles of the analysis.
- Let's assume this becomes the new status quo.
  - Costs of producing the next report on effects of min wage will be very small.
  - Every additional effort will imply improvements on the "state of the art" report (e. g. *dBL*; $\eta(MW), \alpha_1(MW)$)
  - Much easier to have a substantive and normative policy debate (next step:Shiny App!)

Discussion

- Who should work on this:
  - Analytic reviewers of report; Research division within agencies; Study Commissions ("MWSC or MSWC?" [Card and Krueger, 2016])
  - Public policy schools.
  - Think tanks; Bank of knowledge [Clemens and Kremer, 2016].
- Two type of contributions (∼software development):
  - Short term: Within a given time period the model should be taken as given. Less freedom, but direct impact on the policy analysis.
  - Long term: Structural revisions occur in parallel and are incorporated in future cycles of the analysis.
- Let's assume this becomes the new status quo.
  - Costs of producing the next report on effects of min wage will be very small.
  - Every additional effort will imply improvements on the "state of the art" report (e. g. *dBL*; $\eta(MW), \alpha_1(MW)$)
  - Much easier to have a substantive and normative policy debate (next step:Shiny App!)

Discussion

- Who should work on this:
  - Analytic reviewers of report; Research division within agencies; Study Commissions ("MWSC or MSWC?" [Card and Krueger, 2016])
  - Public policy schools.
  - Think tanks; Bank of knowledge [Clemens and Kremer, 2016].
- Two type of contributions ($\sim$software development):
  - Short term: Within a given time period the model should be taken as given. Less freedom, but direct impact on the policy analysis.
  - Long term: Structural revisions occur in parallel and are incorporated in future cycles of the analysis.
- Let's assume this becomes the new status quo.
  - Costs of producing the next report on effects of min wage will be very small.
  - Every additional effort will imply improvements on the "state of the art" report (e. g. *dBL*; $\eta(MW), \alpha_1(MW)$)
  - Much easier to have a substantive and normative policy debate (next step:`Shiny App!`)

Extension To Other Policy Analyses

- Cost Benefit Analysis
- Ex ante economic analysis
- Dissemination

Thank you.

https://fhoces.github.io

Back-up slides

Pierre Azoulay, Christian Fons-Rosen, and Joshua S Graff Zivin. Does science advance one funeral at a time? Technical report, National Bureau of Economic Research, 2015.

David Card and Alan Krueger. Interview with card and krueger. http://davidcard.berkeley.edu/interviews/interview%20with%20Card%20and%20Krueger.pdf, April 2016. (Accessed on 10/26/2016).

Michael A Clemens and Michael Kremer. The new role for the world bank. *The Journal of Economic Perspectives*, 30(1): 53–76, 2016.

Daniel R Feenberg, Ina Ganguli, Patrick Gaule, and Jonathan Gruber. It's good to be first: Order bias in reading and citing nber working papers. Technical report, National Bureau of Economic Research, 2015.

Charles F Manski. *Public policy in an uncertain world: analysis and decisions*. Harvard University Press, 2013.

Sendhil Mullainathan and Eldar Shafir. *Scarcity: Why having too little means so much*. Macmillan, 2013.

BA Nosek, G Alter, GC Banks, D Borsboom, SD Bowman, SJ Breckler, S Buck, CD Chambers, G Chin, G Christensen, et al. Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science (New York, NY)*, 348(6242): 1422, 2015.

Sandra M Nutley, Isabel Walter, and Huw TO Davies. *Using evidence: How research can inform public services*. Policy press, 2007.

Snapshots of DD.

https://rpubs.com/fhoces/dd_cbo_test1

Reader Companion for CBO report on Min Wage (Preliminary Version. Do Not Circulate)

Fernando Hoces de la Guardia + (hopefully) a lot more people

Last edit: 2016-10-16

1 Introduction

The role of policy analysis is to connect research with policy. Because of heavy time constrains, policy analyses are typically ambiguous regarding the details of how the analysis was carried out. This creates three problems: (i) its hard to understand the connection between research and policy, (ii) allows policy makers to cherry pick policy reports, and (iii) hinders systematic improvement and/or automation of parts of the analysis. In this document we demonstrate the use of a reproducible workflow to reduce the ambiguity in policy analysis.

Here we attempt to contribute to the policy discussion of the minimum wage. The minimum wage is a contentious policy issue in the US. Increasing it has positive and negative effects that different policymakers value differently. We aim to add clarity on what those effects are, how much do we know about them, and how those effects vary when elements of the analysis change. We select the most up-to-date, non-partisan, policy analysis of the effects of raising the minimum wage, and build an open-source reproducible analysis on top of it.

In 2014 the Congressional Budget Office published the report titled "The Effects of a Minimum-Wage Increase on Employment and Family Income". The report receive wide attention from key stakeholders and has been used extensible as an input in the debate around the minimum wage[1]. To this date we consider the CBO report to be the best non-partisan estimation of the effects of raising the minimum wage at the federal level. Although there was disagreement among experts around some technical issues, this disagreement has been mainly circumscribed around one of the many inputs used in the analysis, and we can fit the opposing positions in to our framework.

Our purposes are twofold: First, promote the technical discussion around a recurrent policy issue (minimum wage) by making explicit and visible all the components and key assumptions of its most up-to-date official policy analysis. Second, demonstrate how new scientific practices of transparency and reproducibility (T & R) can be applied to policy analysis. We encourage the reader to collaborate in this document and help develop an ever-improving version of the important policy estimates[2] (re)produced here.

Snapshots of DD.

Refe

2 Employment effects

At a general level the effects on employment ($\widehat{\Delta E}$) will be calculated using a more detailed version of the following equation:

$$\widehat{\Delta E} = N \times \eta \times \%\Delta w + \text{Other factors}$$

Where N represents the relevant population, η the elasticity of labor demand, Δw the relevant percentual variation in wages, and the *Other factors* will encapsulate effects on employment through an increase in the aggregate demand.

To describe the methodology behind each of those four components we first describe the data used, the wage variable choose, and the procedure used to forecast the wage and population distribution of 2016 using data from 2013.

2.1 Data, wages, and forecast

To simulate the policy effects we need the distribution of wages and employment under the status quo. From the perspective of 2013, this implies forecasting to 2016 data on employment and wages.

2.1.1 Data

The Current Population Survey (CPS) was used to compute the effects on employment. From the analysis in the section on distributional effects we can deduce that the data corresponds to the Outgoing Rotation Group (ORG). CPS is a monthly cross sectional survey. The same individual is interviewed eight times over a period of 12 months. The interviews take place in the first and last 4 months of that period. By the 4th and 12th interview, individuals are asked detailed information on earnings. The CPS ORG file contains the information on this interviews for a given year. We analyze the data for 2013.

Currently three versions of these data sets can be found online: CPS raw files, ORG NBER and ORG CEPR. The analysis will be performed using the CPER ORG data base.

The weights used in our analysis will be `orgwgt/12`

2.1.1.1 Code to load the data

`R`
`Stata`

Snapshots of DD.

performed using the CPER ORG data base.

The weights used in our analysis will be `orgwgt/12`

2.1.1.1 Code to load the data

R

```
call.cps.org.data <- function(){
  data_use <- "CPER_ORG"

  # Using CEPR ORG data
  if (data_use == "CPER_ORG") {
  # Checking if working directory contains data, download if not.
    if ( !("cepr_org_2013.dta" %in% dir()) ) {
        # create name of file to store data
        tf <- "cepr_org_2013.zip"

        # download the CPS repwgts zipped file to the local computer
        download.file(url =  "http://ceprdata.org/wp-content/cps/data/cepr_org_2013.zip", tf , mode
= "wb" )

        # unzip the file's contents and store the file name within the temporary directory
        fn <- unzip( zipfile = tf , overwrite = T )
    }
    df <- read.dta("cepr_org_2013.dta")
  }

  # Using NBER ORG data
  if (data_use == "NBER_ORG") {
    # Checking if working directory contains data, download if not.
    if ( !("morg13.dta" %in% dir()) ) {
      # Downloading data 53mb
      df <- read.dta("http://www.nber.org/morg/annual/morg13.dta")
    }
    df <- read.dta("morg13.dta")
```

Snapshots of DD.

2.5 Computing effects on employment

Putting all elements together we get:

$$\widehat{\Delta E} = \sum_{g \in \{A,T\}} \left(\widehat{N_g^{final}} \times \widehat{\eta^g_{w \leq MW}} \times \overline{\%\Delta w^g} \right) - \widehat{OF}$$

2.5.1 Code to compute each component

`R`
`Stata`

Components of Elasticities

	Adult	Teen
η_{lit}	-0.03	-0.10
$\eta_{w \leq MW'}$	-0.23	-0.13
F_{adj}	4.50	4.50
$\overline{\%\Delta w}$	13.81	16.65
$\widehat{\eta_{w \leq MW}}$	-0.15	-0.45

Using all the components described above we get $\widehat{\Delta^- E}$ = -478 thousand jobs. The report however computes F^g_{adjs} in a different fashion and gets a value of 4.5 (when computing the values of F^g_{adjs} from the table below - as oppose to using historical values - we get $\widehat{\Delta^- E}$ = -321 thousand jobs).

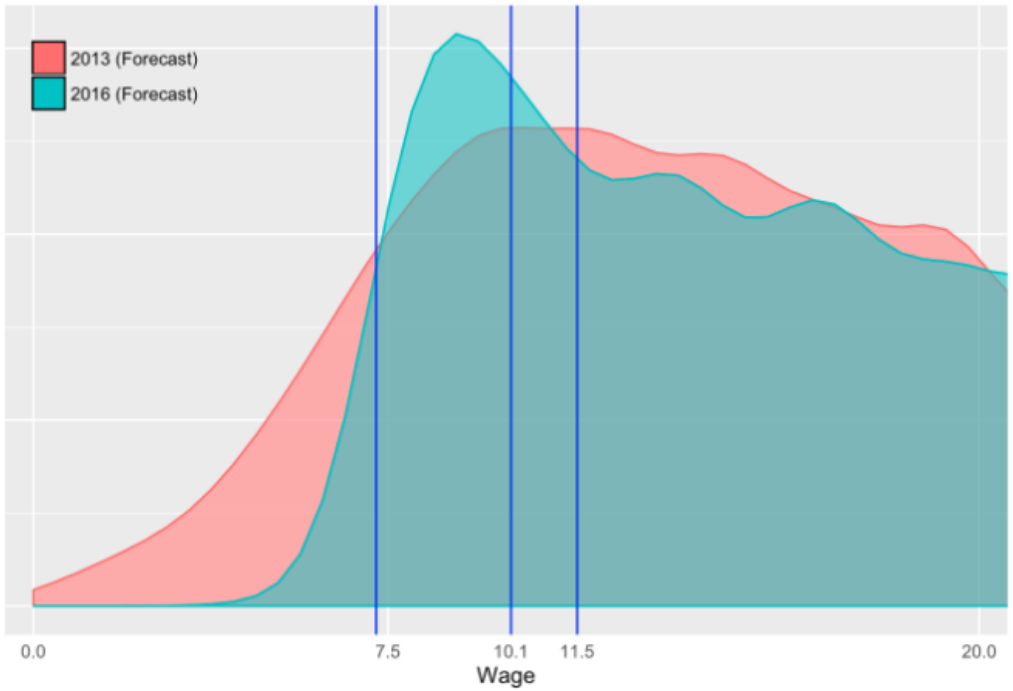
3 Distributional effects

In the first step towards obtaining the policy estimates presented in the introduction we concluded with

Snapshots of DD.

R
Stata

Figure 4: Distribution of wages in 2013 and 2016(forecast)



Comparison of 2013 and 2016 under the status quo

	2013	2016: status quo
Salary workers	122,593,557	129,545,571
Median wage	17.79	20.56

Snapshots of DD.

Final replication output (nothing in the "learn more" button yet)

Policy estimates in CBO report and Replication Results

	Effects/Policy Estimates	Replication
wage gains (billions of $)	31	53.4
wage losses (bns of $)	~5	7.4
Balance losses (bns of $)	~24	43.9
Net effect (bns of $)	2	2
# of Wage gainers (millions)	16.5	23.1/16.9
#of Wage losers (millions)	0.5	0.5

	<1PL	[1PL, 3PL)	[3PL, 6PL)	>6PL
Balance losses (bns of $)	~0.3	~3.4	~3.4	~17
Net effect (bns of $)	5	12	2	-17
Replication loses	-0.4	-6.4	-6.4	-30.8
Replication NE	17.6	14.6	-0.1	-30.1

Learn more

Hinders automation and/or systematic improvements of reports

- Large effort (Res. & PA) to quantify the effects of policies. Effort tends to end with the publication (of paper or report).
- Policy estimates are insensitive to variations in key parameters after publication (interest rate, important prices, cost of technology, etc.).
- We are learning that even the best of analysts, with negligible time constraints, have important biases in their work [Azoulay et al., 2015, Feenberg et al., 2015, Mullainathan and Shafir, 2013].

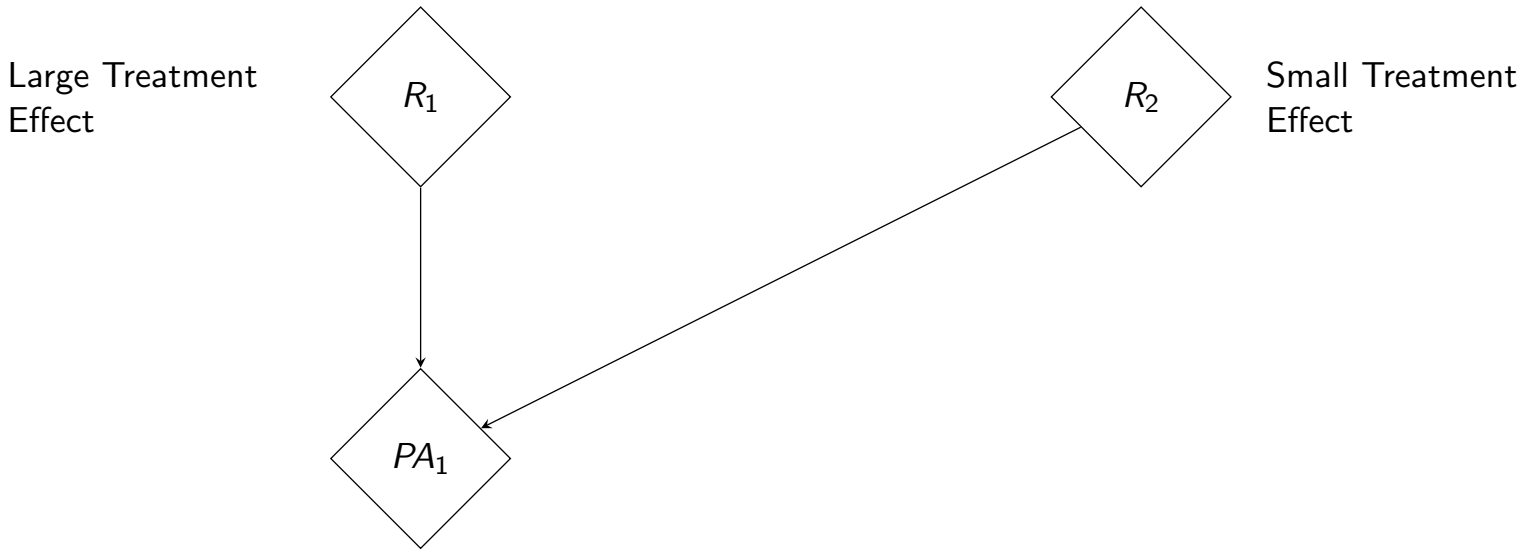
Hinders automation and/or systematic improvements of reports

- Large effort (Res. & PA) to quantify the effects of policies. Effort tends to end with the publication (of paper or report).
- Policy estimates are insensitive to variations in key parameters after publication (interest rate, important prices, cost of technology, etc.).
- We are learning that even the best of analysts, with negligible time constraints, have important biases in their work [Azoulay et al., 2015, Feenberg et al., 2015, Mullainathan and Shafir, 2013].

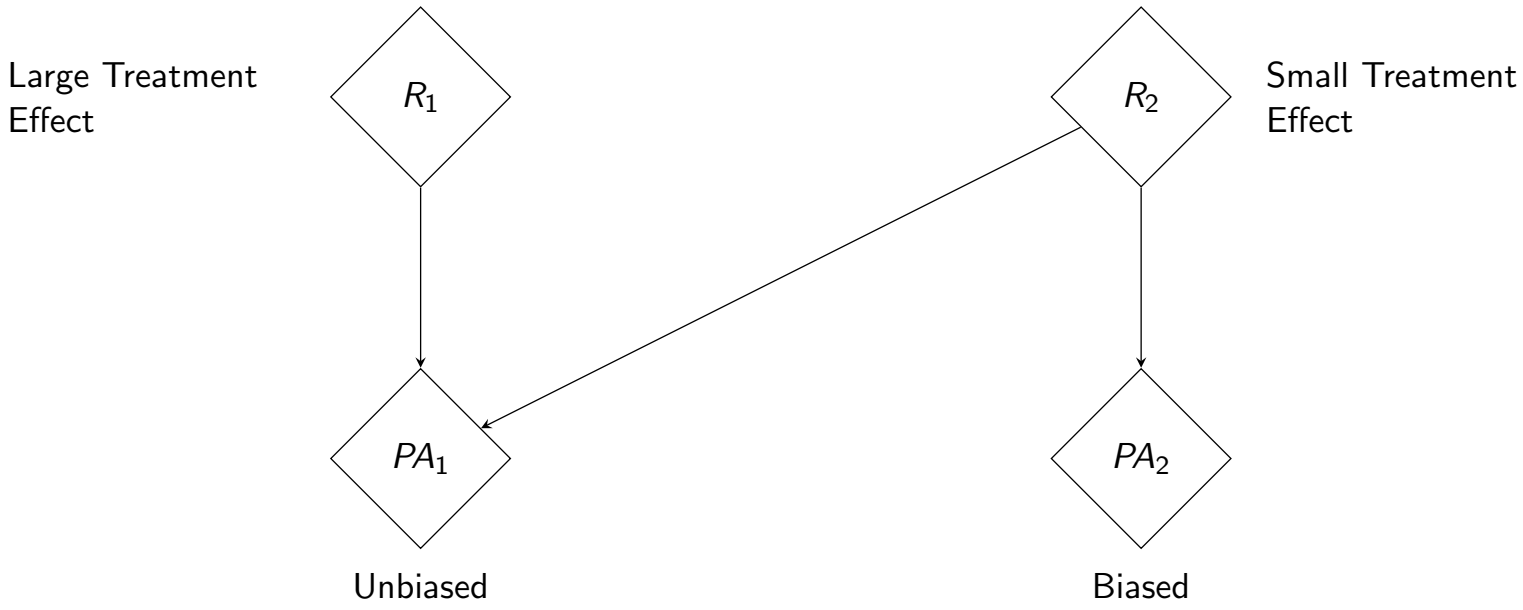
Hinders automation and/or systematic improvements of reports

- Large effort (Res. & PA) to quantify the effects of policies. Effort tends to end with the publication (of paper or report).
- Policy estimates are insensitive to variations in key parameters after publication (interest rate, important prices, cost of technology, etc.).
- We are learning that even the best of analysts, with negligible time constraints, have important biases in their work [Azoulay et al., 2015, Feenberg et al., 2015, Mullainathan and Shafir, 2013].

Threatens the credibility of the policy estimates

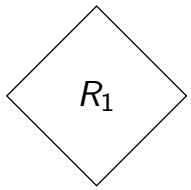


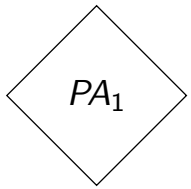
Threatens the credibility of the policy estimates



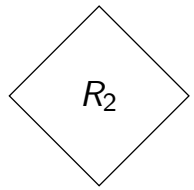
Threatens the credibility of the policy estimates

Large Treatment Effect





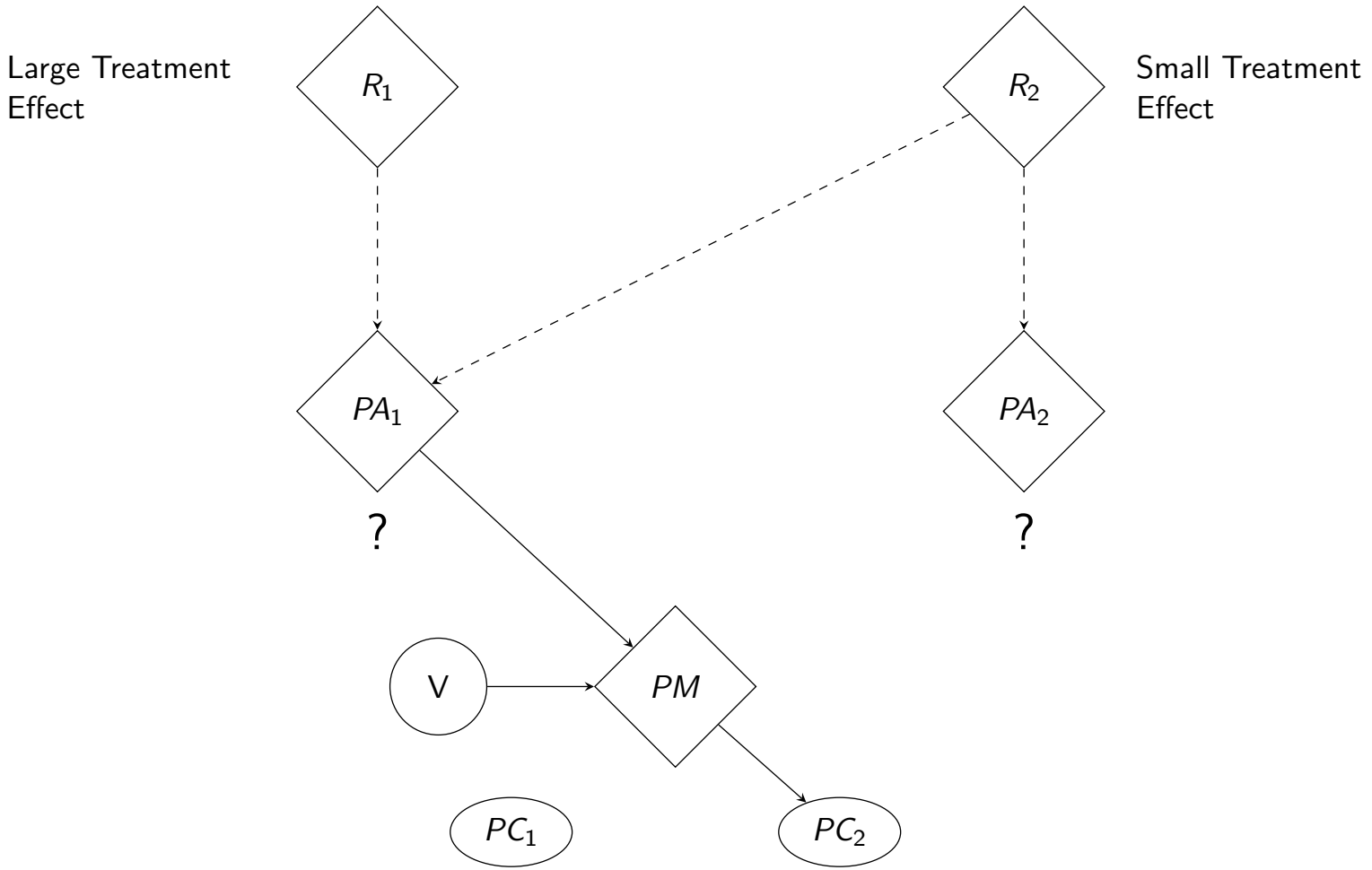
?



Small Treatment Effect

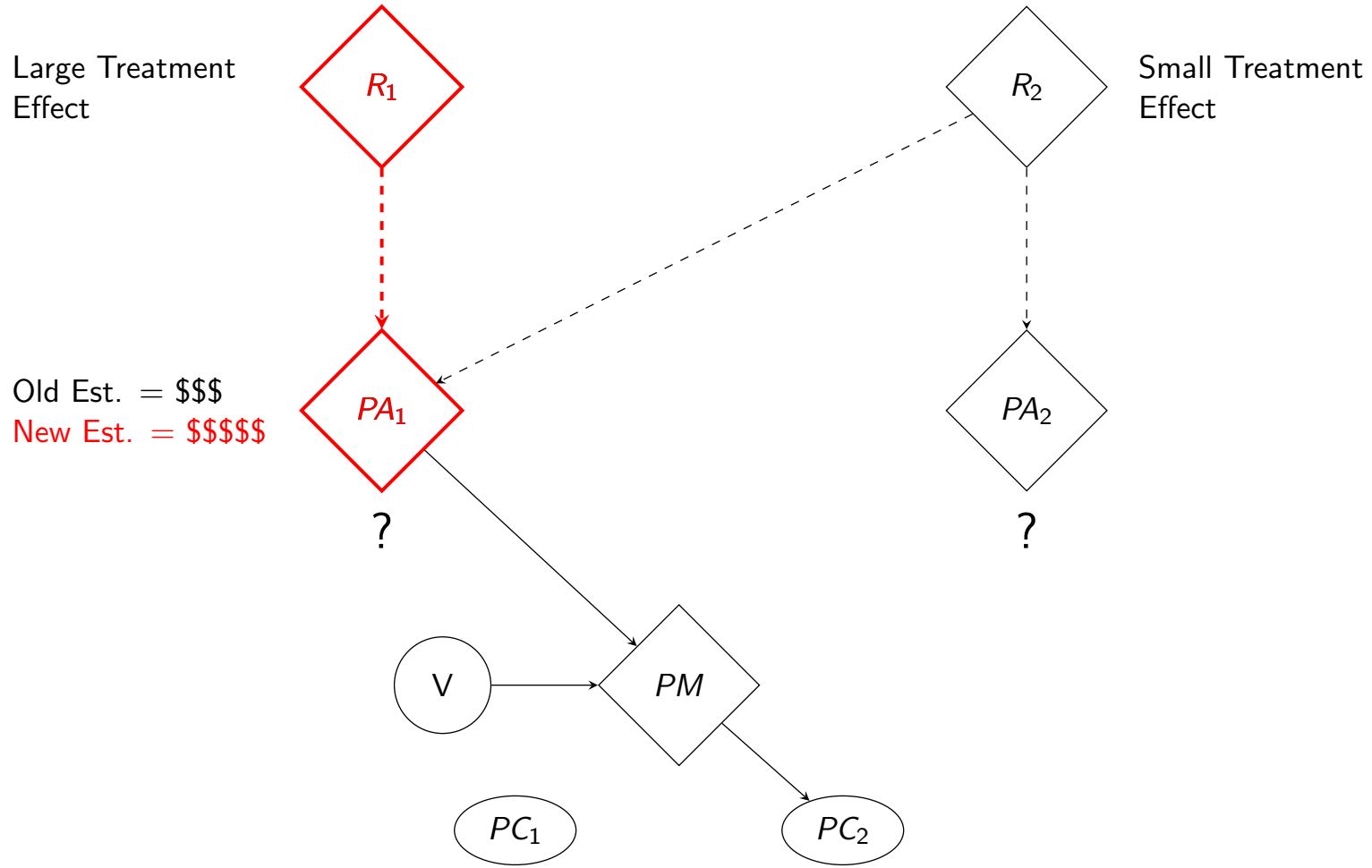
?

Threatens the credibility of the policy estimates



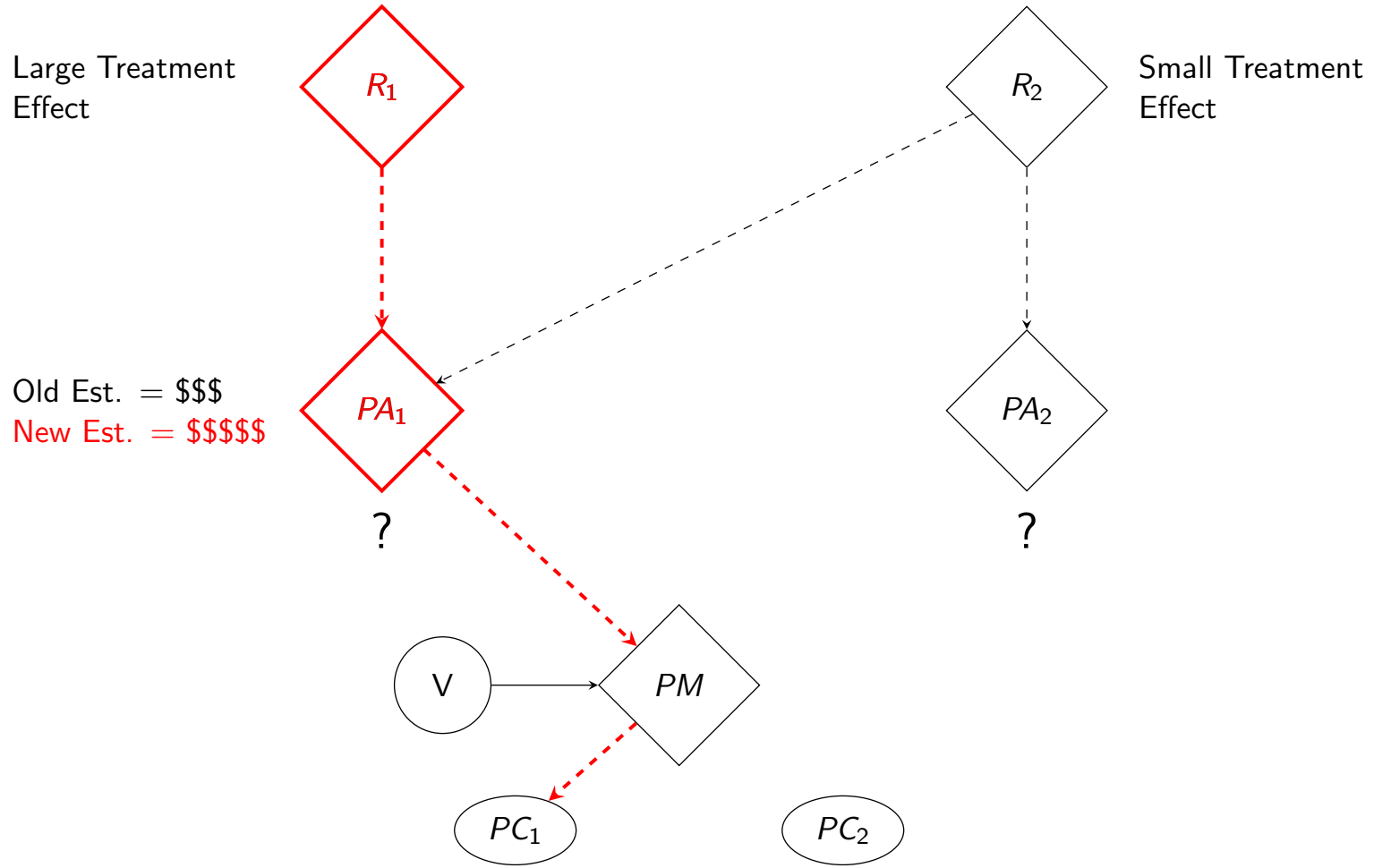
Threatens the credibility of the policy estimates

"When the facts change, _____. What do you do Sir/Madam?"



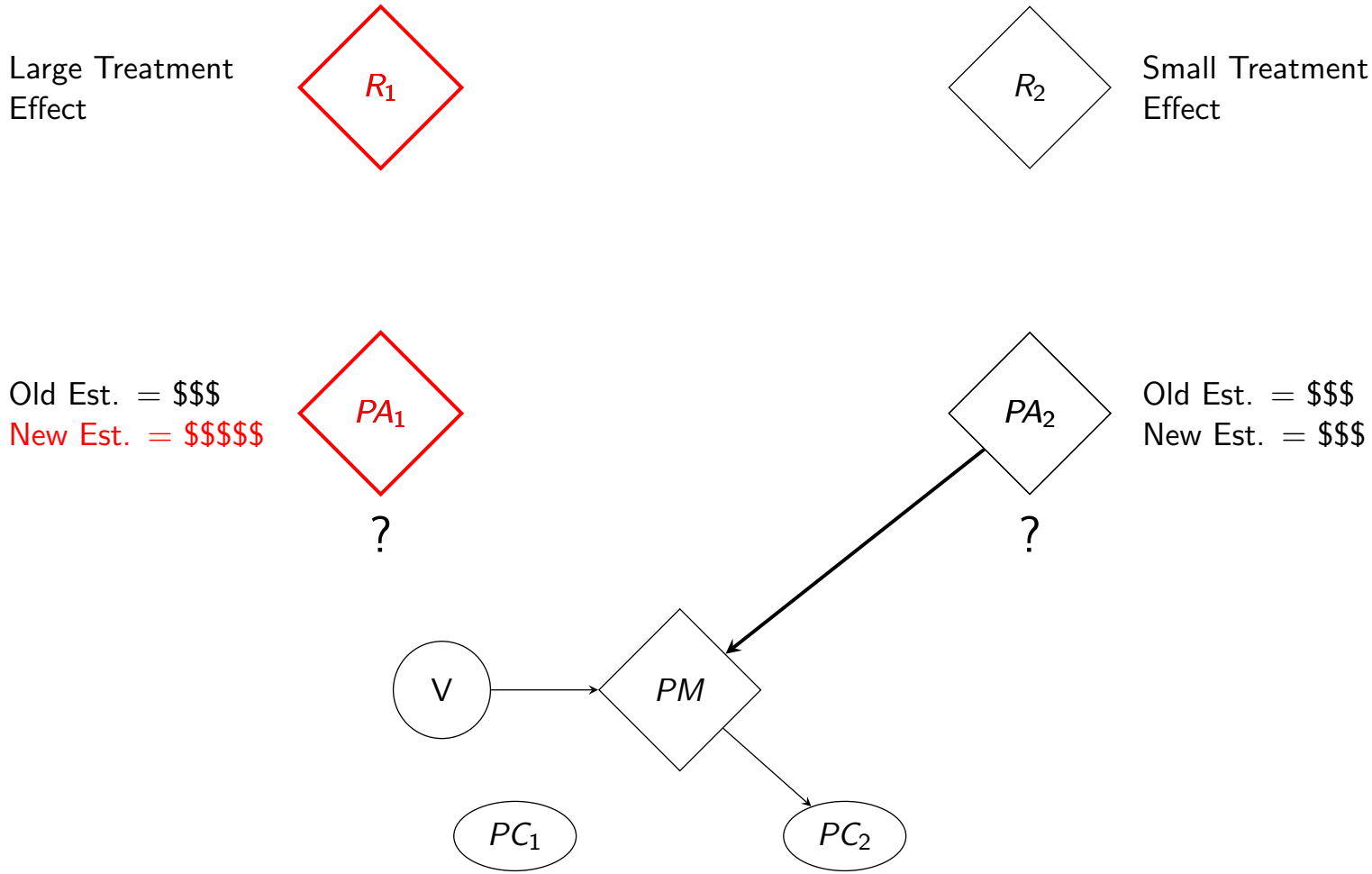
Threatens the credibility of the policy estimates

"When the facts change, I change my mind. What do you do Sir/Madam?" [J.M. Keynes]



Threatens the credibility of the policy estimates

"When the facts change, **I change my facts!**. What do you do Sir/Madam?" [?]

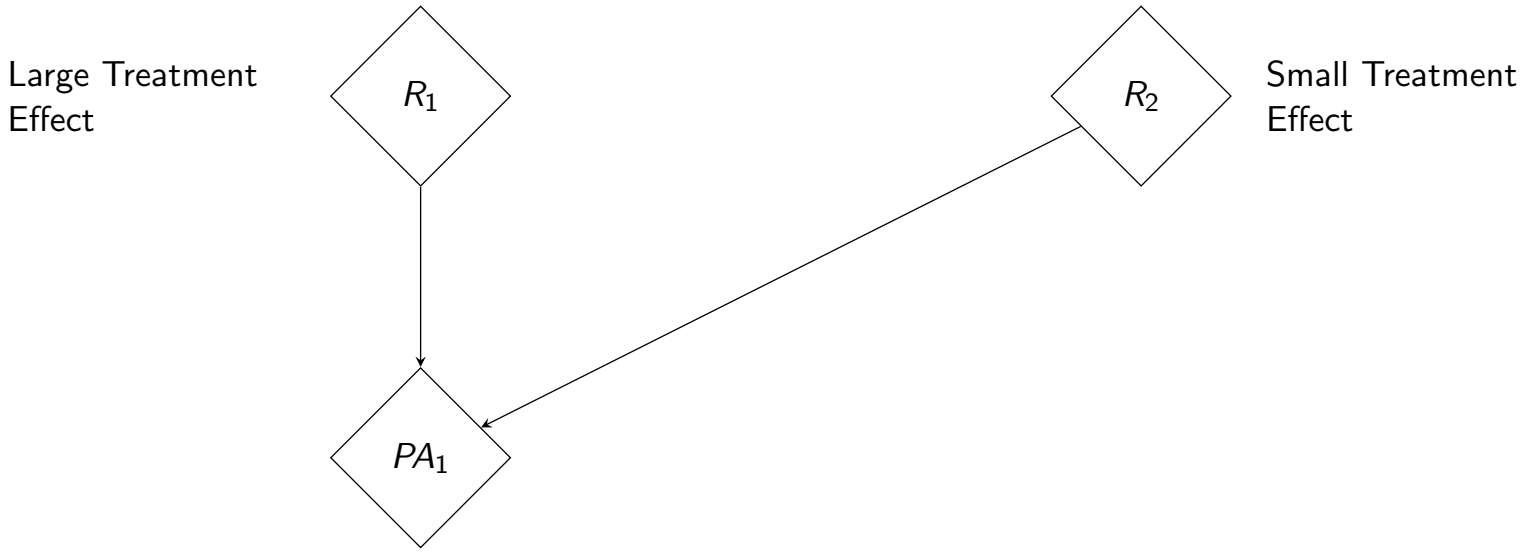


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.

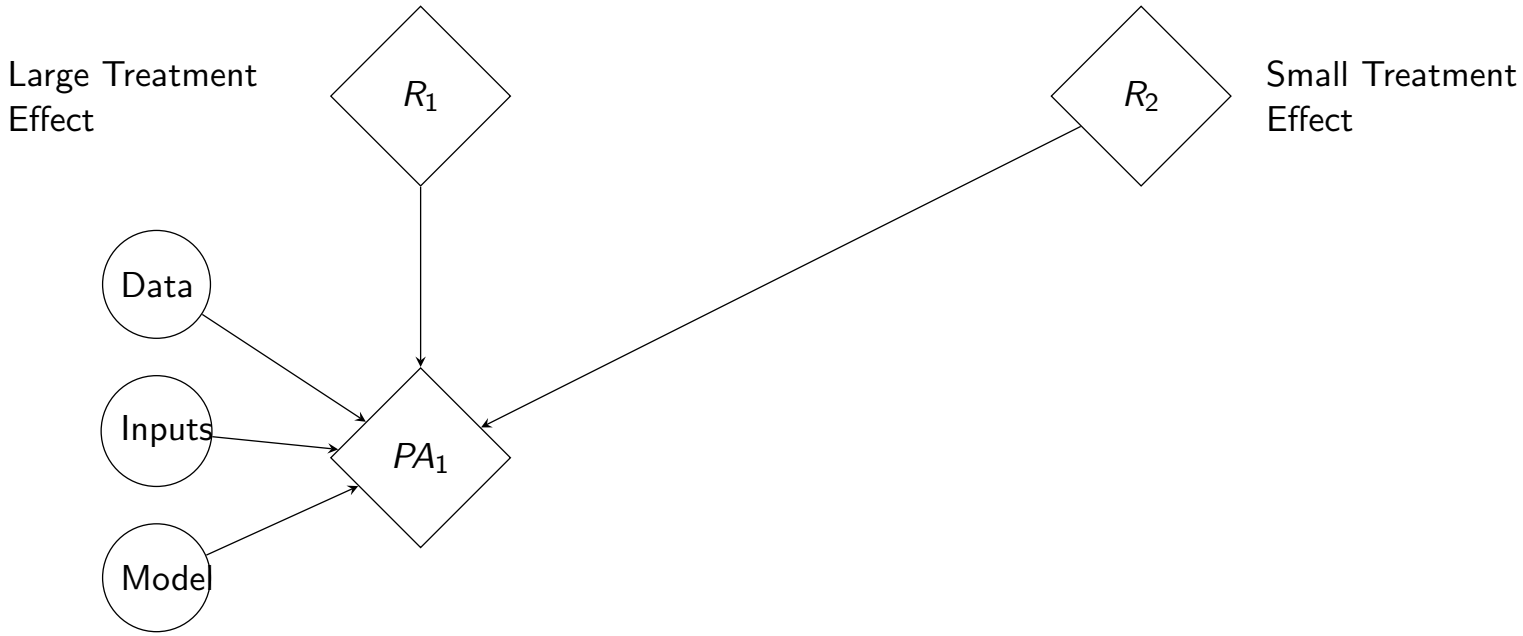


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.

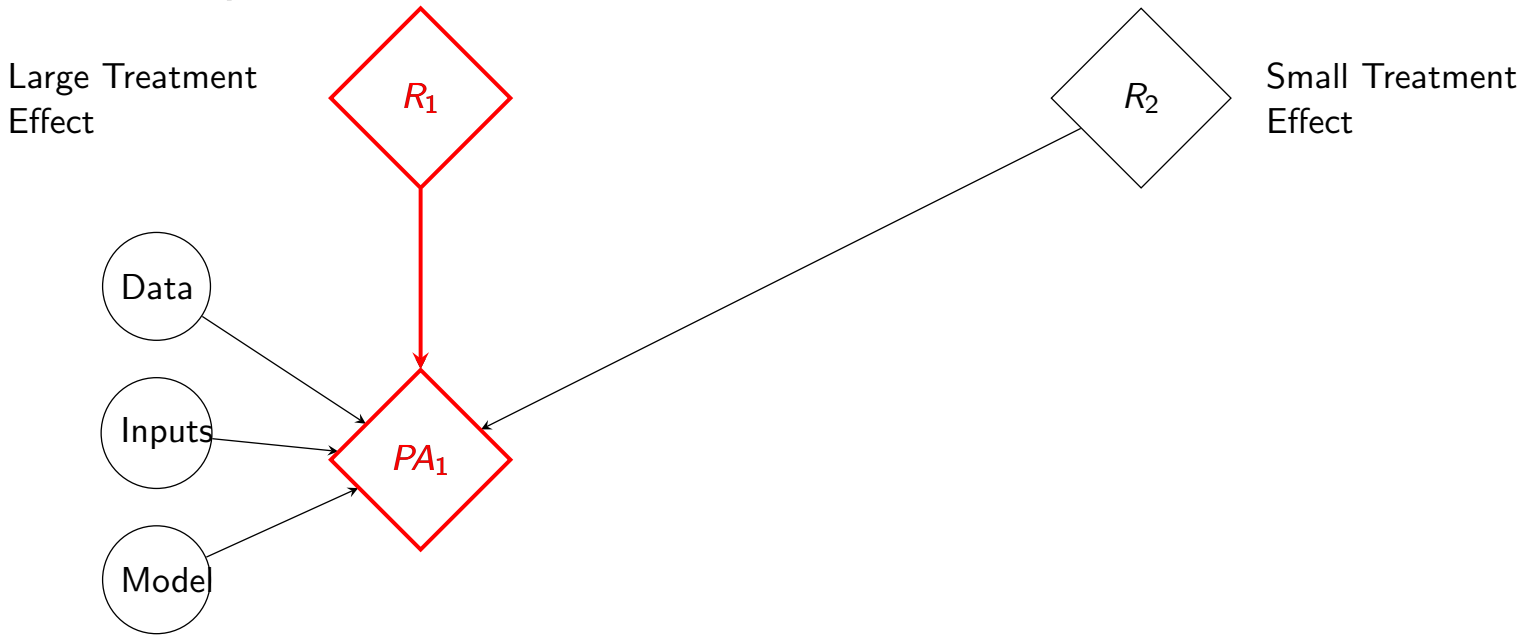


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.

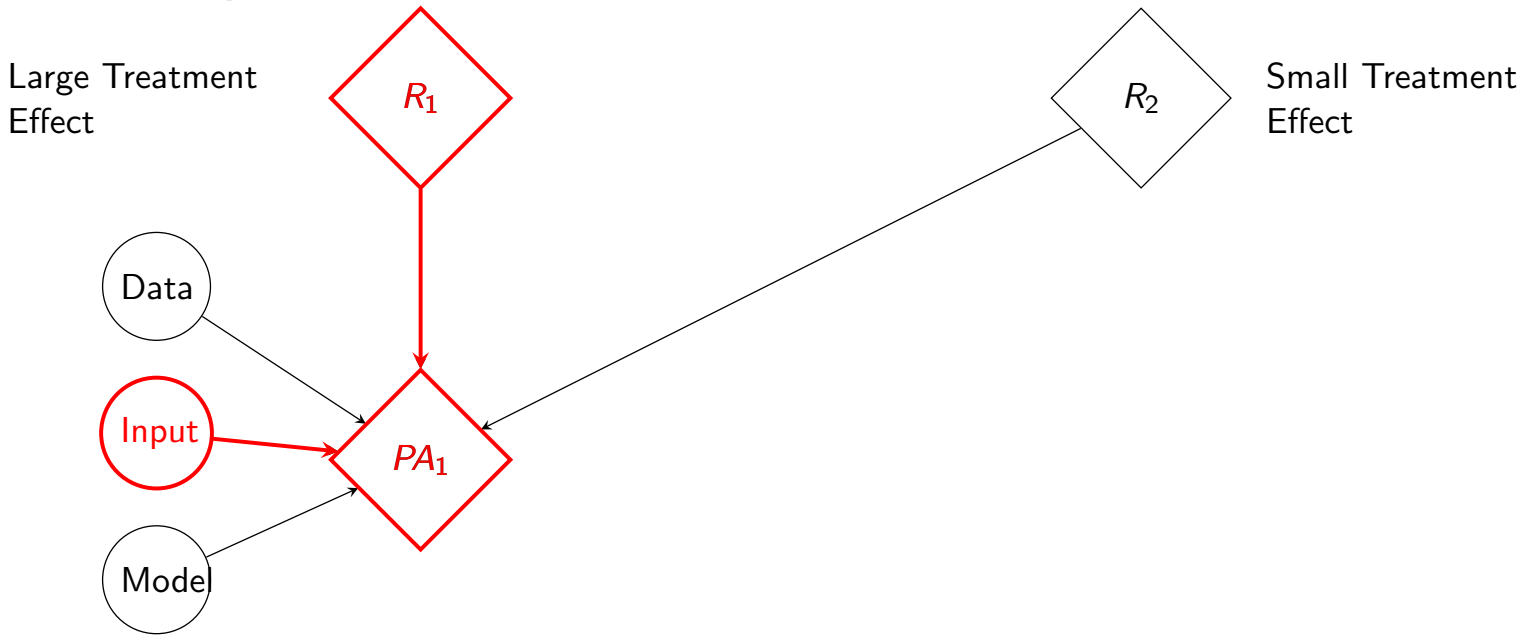


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.

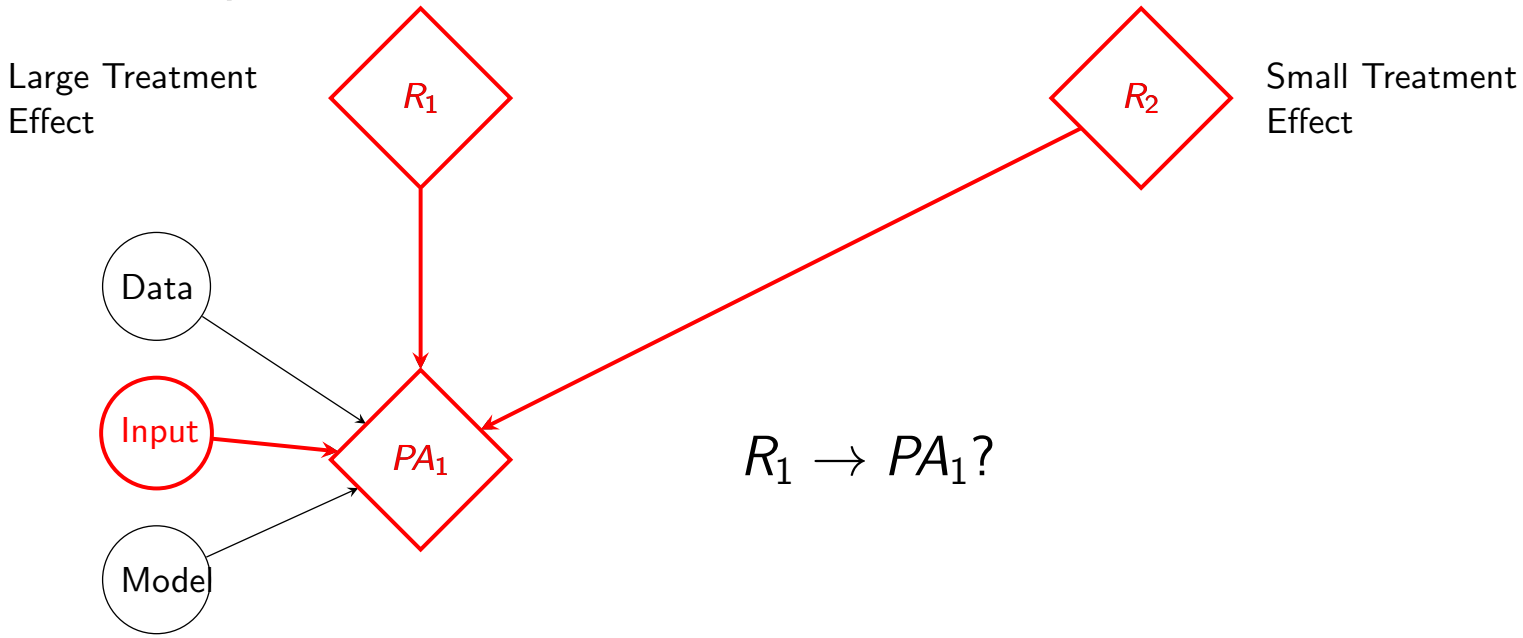


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.

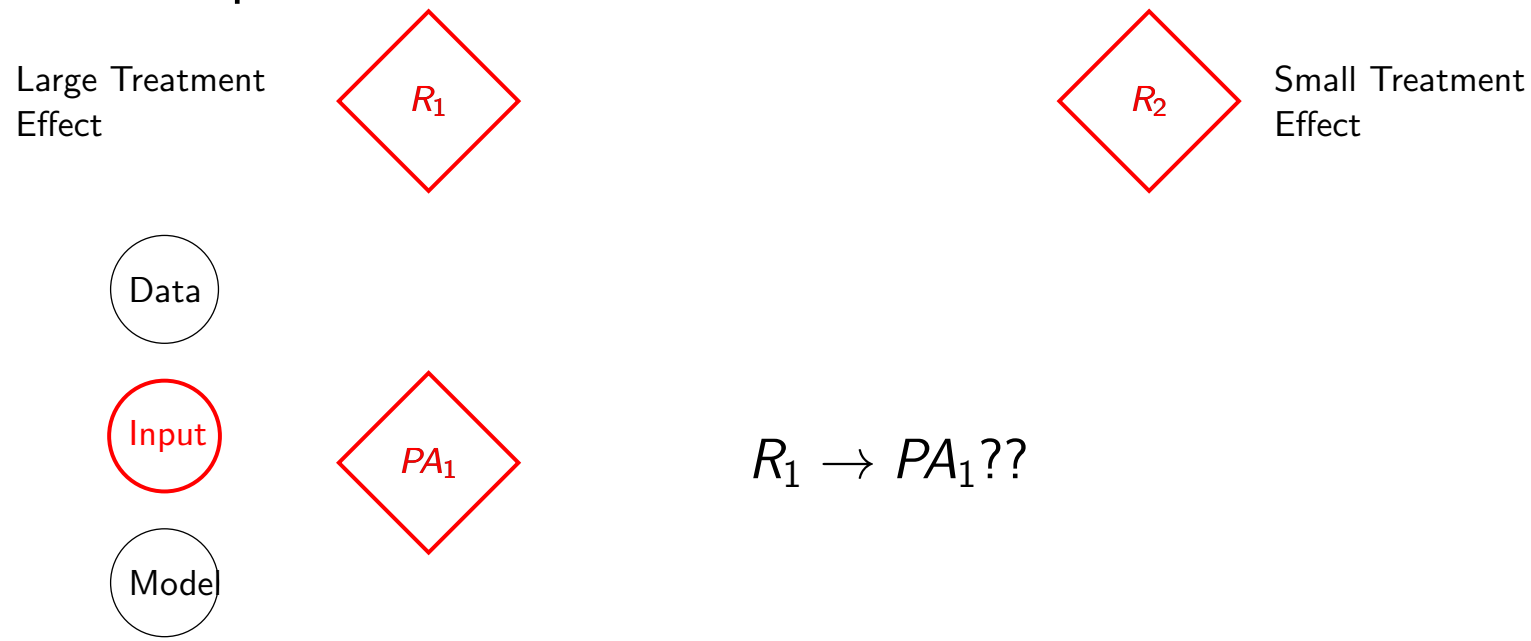


Hard to understand what is the impact of research on policy estimates

What is the impact of R_1 on policy analysis? Can we separate it from a change in inputs? What about R_2?

With the final policy estimate and citations to original research, we can only speculate of impact.

If we see only the final policy estimate, nothing can be said about impact.



Why so little evidence on the role of evidence?

Complex issue

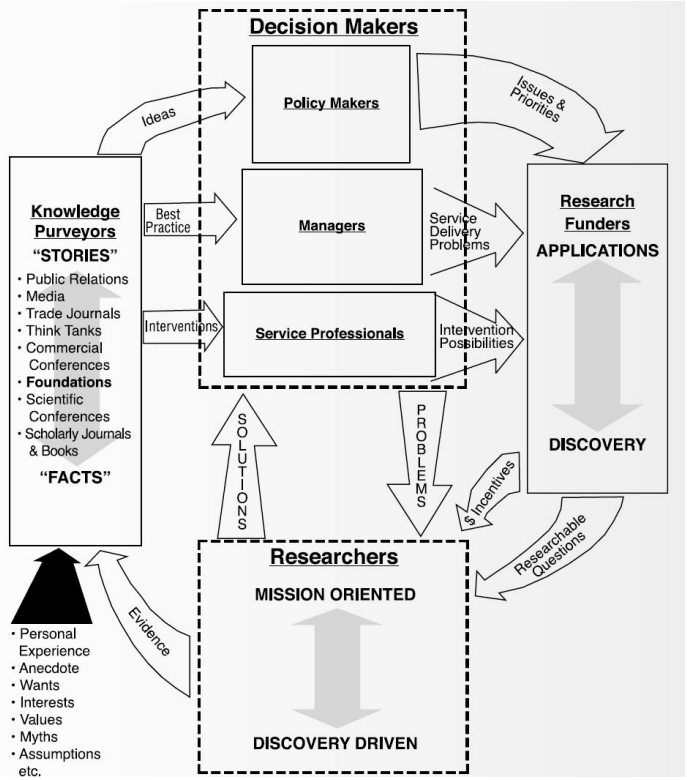


Figure: Reproduced from Nutley et al. [2007] citing CHSRF (2000)

Why so little evidence on the role of evidence?

Complex issue

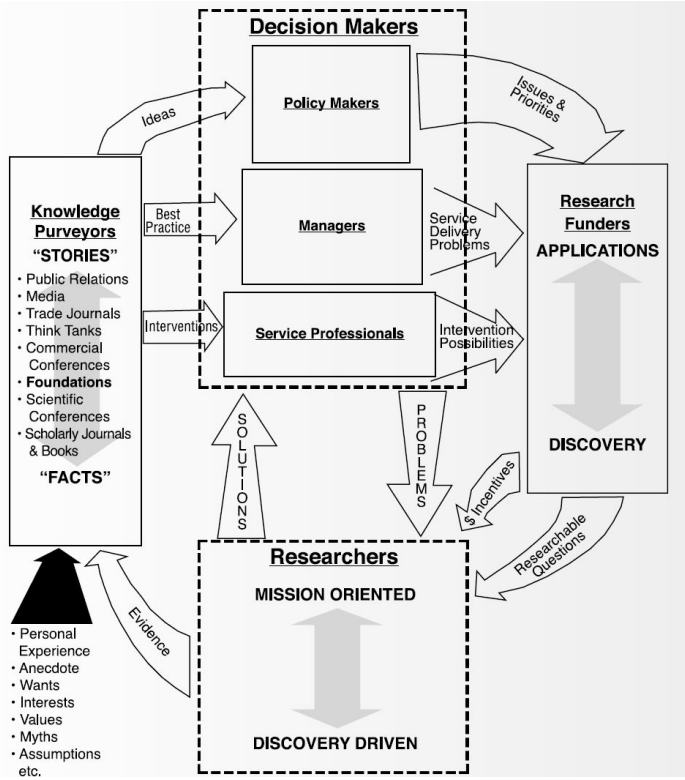
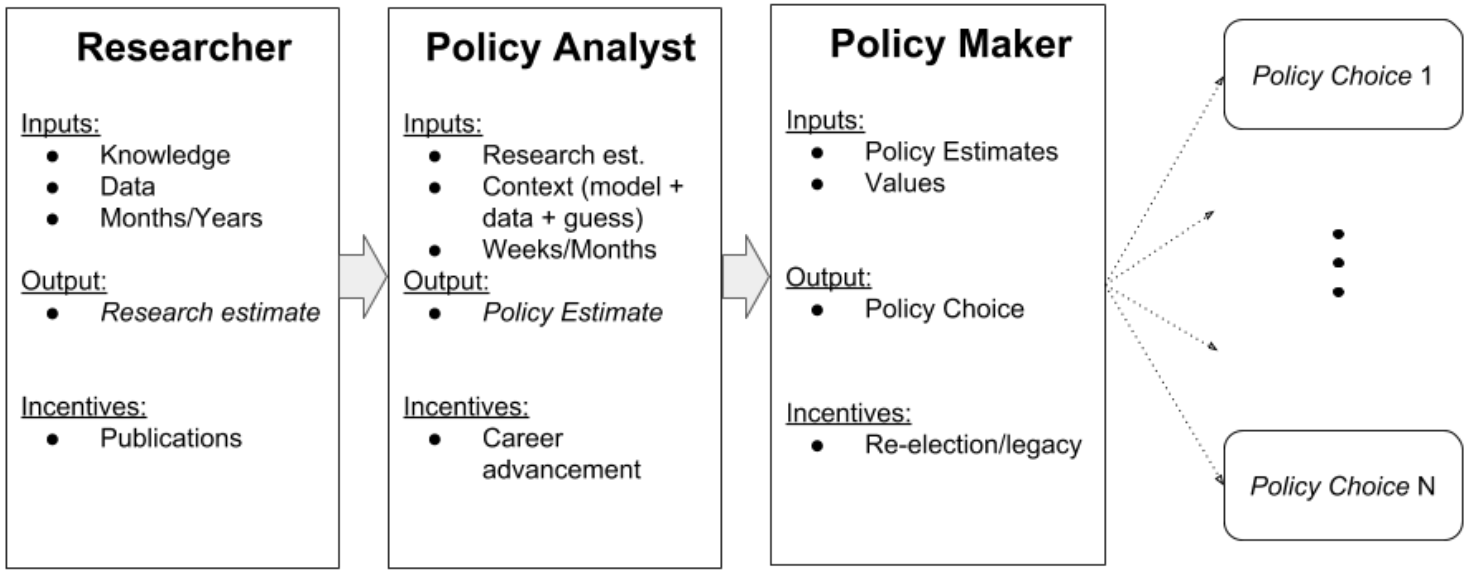


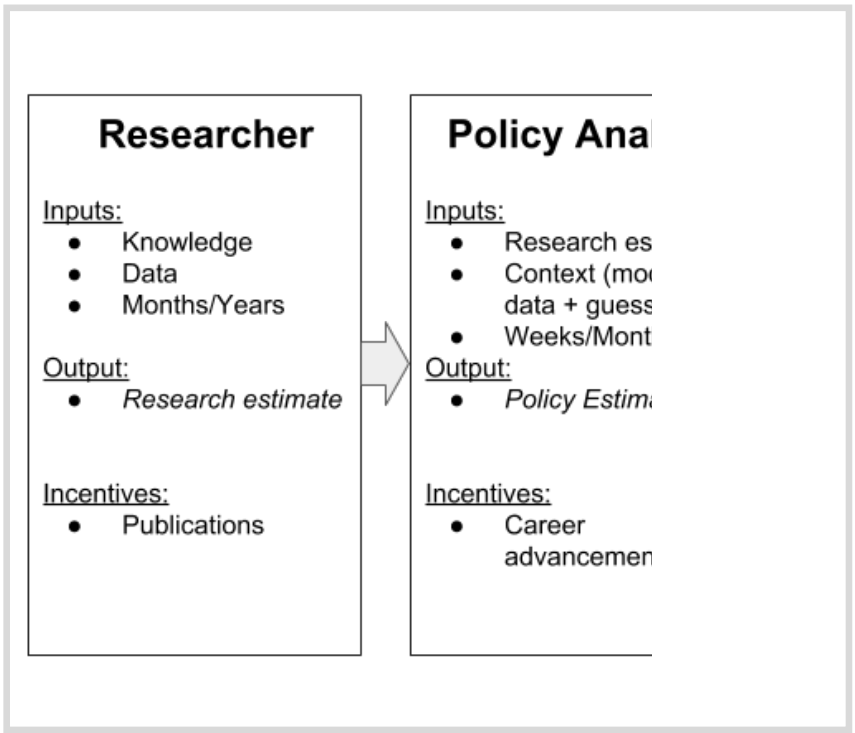
Figure: Reproduced from Nutley et al. [2007] citing CHSRF (2000)

Simpler conceptual model



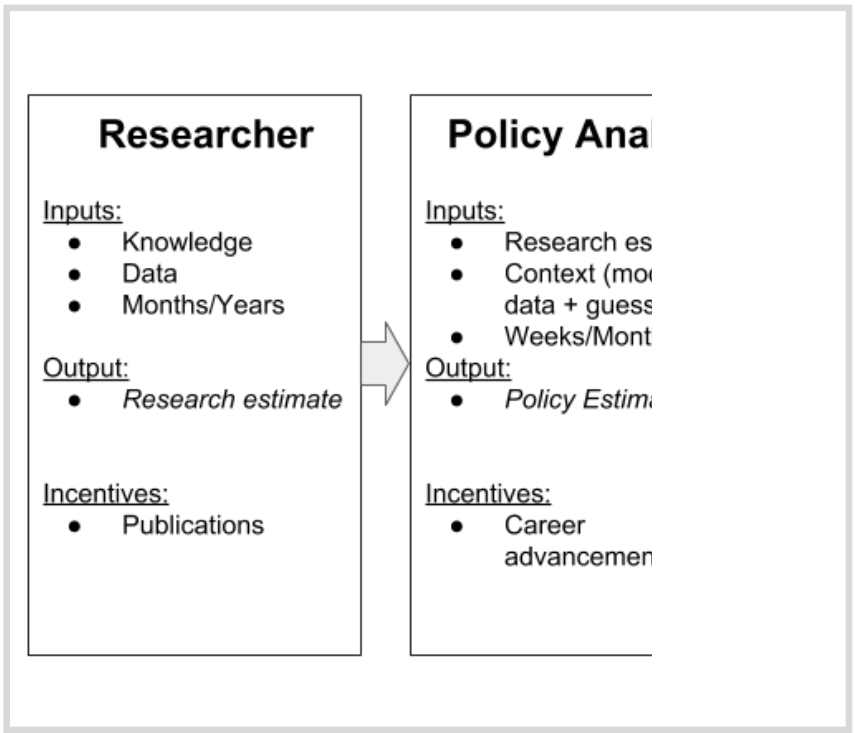
Context: The Role of Evidence in Policy

How does research affect policy analysis?



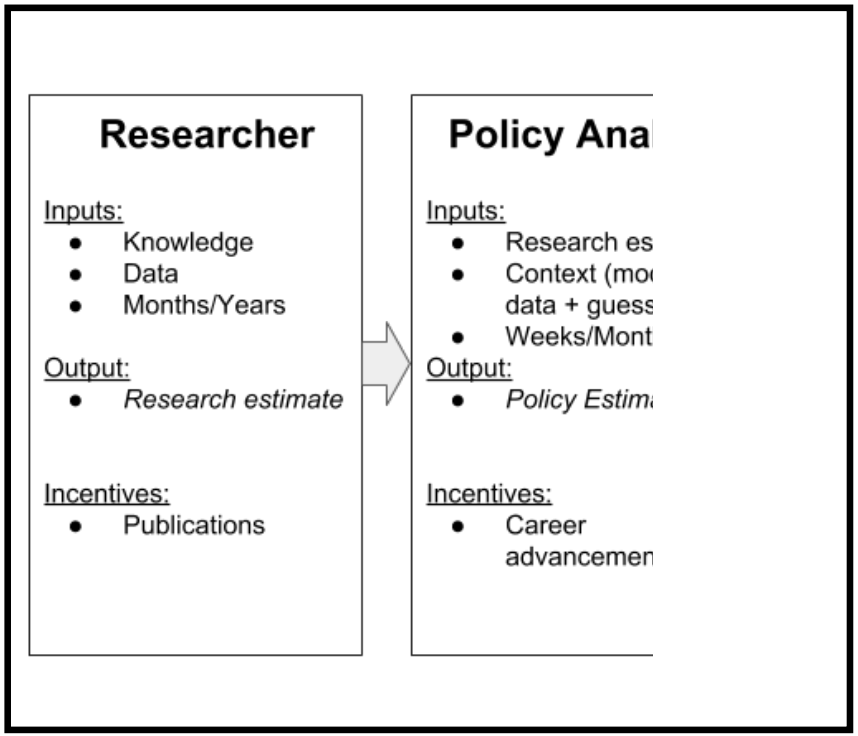
Context: The Role of Evidence in Policy

How does research affect policy analysis?



Context: The Role of Evidence in Policy

How does research affect policy analysis?



Reasons for Selecting the Case Study

- *Scalable*: CBO's reputation: among the most transparent and rigorous policy analysis offices. Lessons from TR that apply to CBO should apply also to most agencies. Additionally the policy issue is widely known which facilitates parallels.
- *Recurrent*: This policy analysis will be conducted again in the future. The case study can be directly used in future calculations.

Reasons for Selecting the Case Study

- *Feasible*: available data, good description of the analysis, and only one policy lever to analyze.
- *Relevant*:

Figure: Google Search Intensity of "Minimum Wage"

Clear connection between sources and inputs

Source	Input
Data	
CPS ORG 2013 (CEPR version)	Number of salary workers in 2013 ($\widehat{N^g_{final}}$ $g \in \{teen, adult\}$); Fraction of workers below the new minimum wage ($P_{\hat{w} \leq MW^1 \mid g}$); Average wage variation for those below the new min wage ($\overline{\%\Delta w^g}$); Non-compliance rate (α_1^g)
CPS ASEC 2012 (CEPR version)	Wages and Non-Wage Income distribution (dF_w, dF_{nw}); Household size (N_h); Hours/weeks worked ($\hat{w}, \hat{h}$)
State level Min. Wage (DOL)	Trends in state min. wage (MW_t^s)
10-year economic forecast (CBO)	Predicted worker growth by 2016 (in 2013) ($\hat{g_N}$); Wage growth in by 2016 ($\hat{g_w}$); Non-wage growth by 2016 ($\hat{g_{nw}}$)
Research	
Elasticity of labor demand for teenagers	$\eta^{lit}_{teen} = -0.1$
Ripple effects	From $R_{lb} = \$8.7$ to $R_{ub} = \$11.5$ with a "ripple" intensity of $R_I = 50\%$
Guess Work	
Extrapolation factor from teenagers to adults	$F_{ex} = 1/3$
Net benefits	$\hat{NB} = \$2 billion$
Adjustment to account for effective wage variation and affected population	$F_{adj} = 4.5$
Aggregate consumption effects on employment	$\hat{OF} = 40,000$ *new jobs*
Distribution of balance loses	$dBL = (1\%, 29\%, 70\%)$ if income $\in [0, 1PL, 6PL, +)$
Fract. of wage loses used to pay wage gains	$F_{subs} = 1$
Job killing process: fraction of jobs	Cut wages in half for twice the number of jobs destroyed

Fully specified model

References

Model	Policy estimate (per quintile)
Predicted household income with and without min wage increase. **Depends on**: $\widehat{N^g_{final}}, P_{\hat{w} \leq MW^1 \mid g}, \overline{\%\Delta w^g}, \alpha^g_1, dF_w, dF_{nw}, N_h, \hat{w}, \hat{h}, MW^s_t, \hat{g^g_N}, \hat{g^g_w}, \hat{g^g_{nw}}, \eta^{lit}_{teen}, R_{lb}, R_{ub}, R_I, F_{ex}, F_{adj}, \hat{OF}$	Average gain in per capita income due to net wage increase. ($\overline{WG_q}$)
Predicted household income with and without min wage increase. **Depends on**: $\widehat{N^g_{final}}, P_{\hat{w} \leq MW^1 \mid g}, \overline{\%\Delta w^g}, \alpha^g_1, dF_w, dF_{nw}, N_h, \hat{w}, \hat{h}, MW^s_t, \hat{g^g_N}, \hat{g^g_w}, \hat{g^g_{nw}}, \eta^{lit}_{teen}, F_{ex}, F_{adj}, \hat{OF}$	Average loss in per capita income due to net wage decrease. ($\overline{WL_q}$)
Distribution of balance loses **Depends on**: $\overline{WG_q}(\cdot), \overline{WL_q}(\cdot), \hat{NB}, F_{subs}, dBL$	Average loss in per capita income to balance wage gains. ($\overline{BL_q}$)