How Transparency and Reproducibility Can Increase Credibility in Policy Analysis: A Case Study of the Minimum Wage Policy Estimates

Fernando Hoces de la Guardia
JMP

First Draft: November 2016
Current Draft: December 2016

Abstract

Observers have pointed out that the analysis of public policies, even when performed by the best non-partisan agencies, often lacks credibility (Manski, 2013). This allows policy makers to cherry pick between reports, or within a specific report, to select estimates that better match their beliefs. For example, in 2014 the Congressional Budget Office (CBO) produced a report on the effects of raising the minimum wage that was cited both by opponents and supporters of the policy, with each side accepting as credible only partial elements of the report. Lack of transparency and reproducibility (TR) in a policy report implies that its credibility relies on the reputation of the authors instead of on a critical appraisal of the analysis.

This paper translates to policy analysis solutions developed to address the lack of credibility in a different setting: the reproducibility crisis in science. I adapt the Transparency and Openness Promotion (TOP) guidelines (Nosek et al, 2015) to the policy analysis setting. The highest standards from the adapted guidelines involve the use of two key tools: dynamic documents that combine all elements of an analysis in one place, and open source version control (git). I implement the highest standards from the adapted guidelines in a case study of the CBO report mentioned above, and present the complete analysis in the form of an open-source dynamic document. In addition to increasing the credibility of the case study analysis, this methodology brings attention to several components of the policy analysis that have been traditionally overlooked in academic research, for example the distribution of the losses used to pay for the increase in wages. Increasing our knowledge in these overlooked areas may prove most valuable to an evidence-based policy debate.

Contents

1 Introduction

In 2014 the Congressional Budget Office (CBO) published a report estimating the effects on employment and income of a potential raise in the federal minimum wage from \$7.25 per hour to \$10.10[1]. The report estimated positive effects due to wage gains, and negative effects due to job loses and income loses from those paying for the wage gains (US Congressional Budget Office, 2014). Supporters of the policy took the positive effects as given, and questioned heavily some of its negative effects (White House, 2014). Conversely opponents of a raise in the minimum wage took the negative effects at face value, while neglecting to acknowledge its benefits (Smith et al, 2014). This example of selective reading of the analysis is common practice in policy, and its supported by the lack of credibility in policy analysis identified by scholars (Manski, 2013). This paper presents a methodological innovation aimed at increasing the credibility of policy analysis.

Manski (2013) describes the practice of policy analysis as one dominated by incredible certitudes. By developing a typology of analyses that presents policy estimates with little or no uncertainty, Manski describes examples of how even the most reputable agencies and research organizations performing policy analysis lack credibility. To address this problem, Manski suggest a menu of methodological improvements to policy analysis. In increasing order of desirability: (1) display standard errors; (2) bound estimated effects; (3) add [policy] decision criteria to the analysis (best).

A underlying theme in Manski's typology is that of characterizing the policy analysis process as a black box, where little is known about how each component of the analysis affects the final results. Using this underlying theme the problem of low credibility can be understood as one of low transparency and reproducibilty (TR). To open the black box this paper draws a connection between Manski's credibility critique and the response to the reproducibility crisis in science. As a complement to Manki's prescriptions, I adapt guidelines and tools from this response to increase TR in policy analysis. This methodological developments are demonstrated in a case study of CBO's analysis on the minimum wage mentioned above.

The deficit in TR in policy analysis harms public policy for three reasons. First, it makes it difficult to understand precisely how research affects the estimations produced by policy analysis (policy estimates hereafter). Research generates evidence that feeds into policy analysis. But little is known about how the emergence of new evidence might affect those policy estimates (Vivalt, 2015; Nutley et al, 2007). If we were to completely understand how the current research is used in policy analysis, we could accurately assess the potential effect of new research on policy estimates. With this information, research resources could be allocated based on where the value of additional information is the highest, and researchers could identify the biggest gaps in knowledge related to specific policies (Snilstveit et al, 2013).

Second, lack of TR prevents automation and/or systematic improvements of reports. Translating evidence from general research into specific policy estimates is not a easy task. It requires modeling how multiple agents will react

[1]The report explored two policy alternatives: \$10.10 indexed to inflation, and \$9 without indexation. For simplicity only the 10.10 option is described, but the analysis of the paper applies to both.

to a specific policy, assumptions regarding generalizability, estimation of key socio-demographic parameters for the target population, and other contextual elements (like relative prices of inputs or rates to discount effects over time). All these processes are usually encapsulated under the task of generating a micro-simulation study. Micro-simulation studies require a large amount of highly skilled labor, which tends to be expensive, and involve a large number of somewhat arbitrary decisions that need to be made under heavy time constraints. With high levels of TR, redoing this analysis or performing similar ones can be substantially less expensive. With the current lack of TR in policy analysis, such savings cannot be realized. Moreover there is no reason to think that the arbitrary assumptions chosen for one report will be consistent with ones chosen for future versions of similar reports.

Third, lack of TR implies that trust on policy analysis is based on reputation as opposed to sound reasoning. When discussing the Congressional Budget Office, Manski comments "[...]I worry that someday sooner or later the existing social contract to take CBO scores at face value will break down. Conventional certitudes that lack foundation cannot last indefinitely." (Manski, 2013, page 20) Low credibility or TR in non-partisan estimates makes it easier for different parties to cherry-pick their facts from less neutral policy organizations. In turn, overt cherry picking provides a fertile ground for the surge of demagoguery and a general disregard for a scientific approach to policy.

This papers proposes a systematic approach to increase TR in policy analysis. This approach defines the three study objectives:

1. Translate guidelines developed to increase TR in science to the policy analysis setting. After describing the well documented replication crisis and its proposed solutions, common elements are identified between the low TR of policy analysis and science. This comparison also highlights the role of tools, like dynamic documents and open source version control (git), to achieve the highest levels of TR.

2. The drafted guidelines are implemented in the case study of minimum wage policy analysis. First a TR assessment is briefly discussed for the current state of the report. Second, I implement the highest standards from the adapted guidelines in a case study of the CBO report mentioned above, and present the complete analysis in the form of an open-source dynamic document[2]. This document is compared to the original report to demonstrate the gains in TR.

3. In addition to increasing the credibility of the case study analysis, this methodology brings attention to several components of the policy analysis that have been traditionally overlooked in academic research. The fully reproducible report allows for a comprehensive sensitivity analysis showing how several components of the analysis are as important, or more, than the much investigated elasticity of labor demand. For example the distribution of the losses used to pay for the increase in wages. Increasing our knowledge in these overlooked areas may prove most valuable to an evidence-based policy debate.

[2]The first version will be deployed on github on December 10. For later versions of the dynamic document I intend to use the Open Science Framework.

The background section provides information about the reproducibility crisis and about the case study. The methods section describes the adoption of guidelines, the replication of the case study using dynamic documents to achieve the highest standards of those guidelines, and the sensitivity analysis. The results section illustrates differences between the original report and reproduced version using the highest standards of TR. In the same section a sensitivity analysis shows how the output of the report changes when its components vary. The final section discusses the limitations and next steps for increasing TR in policy analysis.

2 Background

2.1 Description of the reproducibility crisis in science and its response

More than four decades ago social scientists from different fields identified how the lack of TR threatened the validity of scientific results. By not disclosing most of the research decisions made in studies, two problems emerge: First, Rosenthal (1979) identified the "File Drawer Problem", currently known as Publication Bias, where only the studies with strong results tend to get published. Second, Leamer (1983) suggests that the inability to observe all of the model specifications tested by the researcher invalidates most of the conclusions presented in empirical work. Both problems could only be solved by opening up the entire scientific process, which means providing access to the original raw data and detailed instructions for all of the analysis conducted as part of the study. Unfortunately these critiques were largely ignored by the scientific community for years.

Publication bias and specification search harm science, and they are the best documented consequences of lacking TR (Jager and Leek, 2013; Ioannidis, 2005; Franco et al, 2014; Simonsohn et al, 2014). The more general problem is how sharing only a small fraction of knowledge generated in a study is in direct contradiction with the most basic goals of science (Merton, 1973). Policy analysis and science have different goals but share key similarities. Both agree with using the best knowledge available to the advancement of society, and both use knowledge as their main input. Science uses knowledge as an input to generate more knowledge. Policy analysis uses knowledge to generate condensed information designed to brief policy makers. Hence, the lack of TR harms both Science and Policy Analysis, but only the consequences of the former are currently documented.

In the last decade evidence of lack of TR in the social and bio-medical sciences appears to have accumulated to a critical point (Collaboration et al, 2015; Nosek et al, 2015; Miguel et al, 2014).

One emblematic case where TR could have saved billions of dollars is the initial studies that analyzed the effects of the drug Tamiflu, which is used to treat seasonal flu. Initial evidence suggested that Tamiflu could ameliorate several harms related to the flu, and the governments of the US and the UK bought more than $20 billion in stockpiles. However once more researchers were able to access the original data and other studies that had been ignored in the initial review of the literature, all the benefits from Tamiflu disappeared. The

British Medical Journal (BMJ) has emphasized that these financial resources would have been saved if protocols for higher TR had been in place (Abbasi, 2014).

This is one example for one scientific discipline, but across fields different types of initiatives are taking place. The BMJ and other top medical journals have adopted new standards for publication, which has increased TR substantially. In the social sciences, initiatives like the Center For Open Sciences (COS) and the Berkeley Initiative for Transparency in the Social Sciences (BITSS) have created tools, conducted research and gathered resources to promote TR. Efforts in this direction have also found strong support from the fast-growing field of Data Science. Both in the public and private sector, this field has made a hallmark of their work the promotion of open access to code, analysis and data (Peng, 2011).

Miguel et al (2014) argues for three guiding norms to promote transparency in social sciences: (i) disclosure of key details involved in the analysis and collection of the data; (ii) registration of pre-analysis plans that contain information on the outcome variable, independent variable(s) of interest, model specifications and other analytical choices before the data is collected; and (iii) open access to data, code, and additional documentation. Nosek et al (2015) operationalize these norms in to a set of guidelines that identify different levels of compliance. They separate the different dimentions of openness in to eight standards. Each standard is presented in a way that allows to qualify different levels of compliance. This allows to understand the challenges of TR in a continuum from not reproducible at all to full transparency and reproducibility (Peng, 2011). Christensen and Soderberg (2015) provide a detailed manual on how to achieve these standards.

Table **??** presents a comparison of problems and solutions associated with lack of TR, between 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)

Of the three norms described in Miguel et al (2014) to increase TR in Science, two can be translated directly to policy analysis: disclosure of key details and open access to all data and materials. Applying such norms in policy analysis should be seen as one more step of the ongoing improvements towards a more open government (White House, 2013). This step would involve adding a *open methods* component to the growing governmental initiatives of *open data* around the world (Gray, 2014). The translation of the third norm regarding preregistration is not as direct.

An important difference on the proposed norms between the TR practices in Science relative to Policy Analysis has to the do with the nature of the output produced in each case. The output for which research is traditionally judged are peer reviewed articles, and they reflect and advancement in the state of knowledge in some specific domain. To avoid the file drawer problem and publication bias the norm of preregistration is proposed to encourage the display of all the scientific analyses, as oppose to only the ones that result in "interesting" results (strong effect size, statistically significant). The output to-judge in Policy Analysis are policy reports. They represent the best estimation available on a specific policy issue chosen by an official agency, and are produced under a much more stringent time constrain. Preregistration is not feasible in this environment. Policy reports are not statements about the truth to be advanced, but the best possible answer to a policy-related question given important resource constrains. This apparent disadvantage can be turned into turned in to a strength: policy reports could, in principle, become "living documents" subject to continuous updating and scrutiny from analysts, academics and the general public. Currently the practice of public comment to legislation/regulation in the US represent a step in this direction.

2.2 Description of the case study

In 2014 the Congressional Budget Office published a report estimating the effects on employment and income of a potential raise in the federal minimum wage, from $7.25 per hour to $10.10. The minimum wage was raised for the last time in 2007 with no adjustments for inflation, so had decrease in real value since. The new proposal involve indexing the new minimum wage to inflation[3]

The report estimated positive effects due to wage gains, and negative effects due to job loses and income loses from those paying for the wage gains (US Congressional Budget Office, 2014). The total wage gain was estimated to be $ 31 billions for 16.5 million workers. The number of jobs lost was estimated to be around 500,000. The net distributional effects where the following: $5billion net total gains for households below the poverty line (PL); +$12billion for households between one and three PL; +$2billion between three and six PL; and net total loss of $17 billion for household with incomes above six PL.

The research on effects of minimum wage on teenage employment is well developed in the US. This effects are measured by estimating the elasticity of labor demand for this population. A intense debate on this effects has driven the research agenda for more than two decades (Card and Krueger, 2015; Neumark and Wascher, 2008; Dube et al, 2010; Clemens, 2015). The findings can be grouped into two schools: one documents large effects on employment with

[3]The $9 option described in the previous footnote was not index to inflation. All the effect estimated by CBO where to the year 2016 so indexation did not made a substantive difference.

elasticity estimates concentrated around -0.1, and the other reports small effects on employment with estimates concentrated around -0.01[4].

The case study on the minimum wage policy analyis was chosen based on four criteria: relevancy, generalizability, recurrence and feasibility.

First the report was clearly influential in the policy debate. It was cited by proponents and opponents of the raise, and it was feature prominently in the news and editorials of that period. As an illustration, figure 1 shows how the publication of this report coincides with the highest search intensity in google for term "minimum wage" in the US.

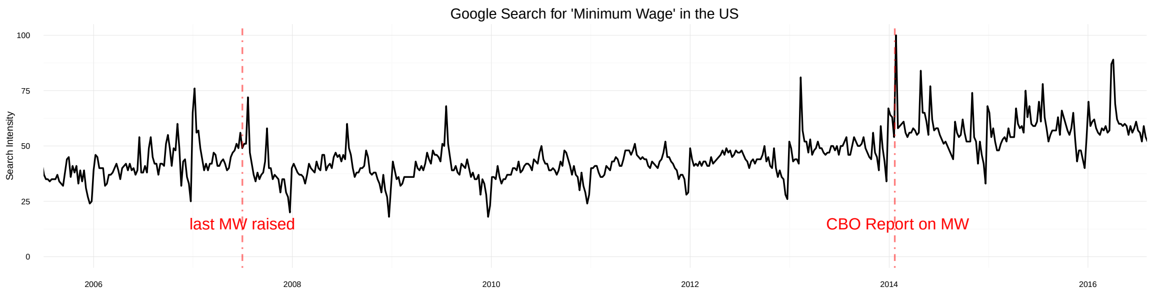


Figure 1: Google Search Intensity of "Minimum Wage"

Second, CBO is among the most transparent and rigorous office of policy analysis. The protocols and tools discussed below should be understood as one additional layer of TR, to the best practices already present in the report. Lessons from TR that apply to the CBO report should apply also to the work of most official agencies. Additionally the policy issue is widely known which facilitates extrapolation to other policy analyses.

Third, the discussion around the minimum wage in the US is notoriously recurrent. This makes highly likely that a similar policy analysis will be conducted again in the future. The case study can be directly used in future calculations.

Finally, it was feasible. All the data was publicly available, the report describes in great detail its analysis, and had only one policy lever to analyze (the minimum wage level).

3 Methodology

To address the low credibility critique of policy analysis (Manski, 2013), the general methodological suggestion of this paper is to increase its transparency and reproducibility (TR).

Figure 2 represents a policy analysis that is completely transparent and fully reproducible. The connection between each source of information (data, research, guess work) and the final policy estimates is perfectly known. The arrows connecting each component reflect the knowledge of how all the pieces interact together. For example, with full TR it can be known exactly how the

[4]This academic debate transcends the minimum wage policy debate in at least two dimensions: in economics has come to challenge the predictive power of the most basic models taught in introductory classes, and in empirical research in general the debate is contemporaneous to a debate between the relative importance of empirical strategies vs theoretical predictions.

data was used to generate each input, and how those were used in the model to compute the final policy estimates. Reproducibility, although not explicit in the diagram, is the key feature that allows users (e.g. analysts, researchers) to modify/test any component and asses its effect on the final policy estimates.

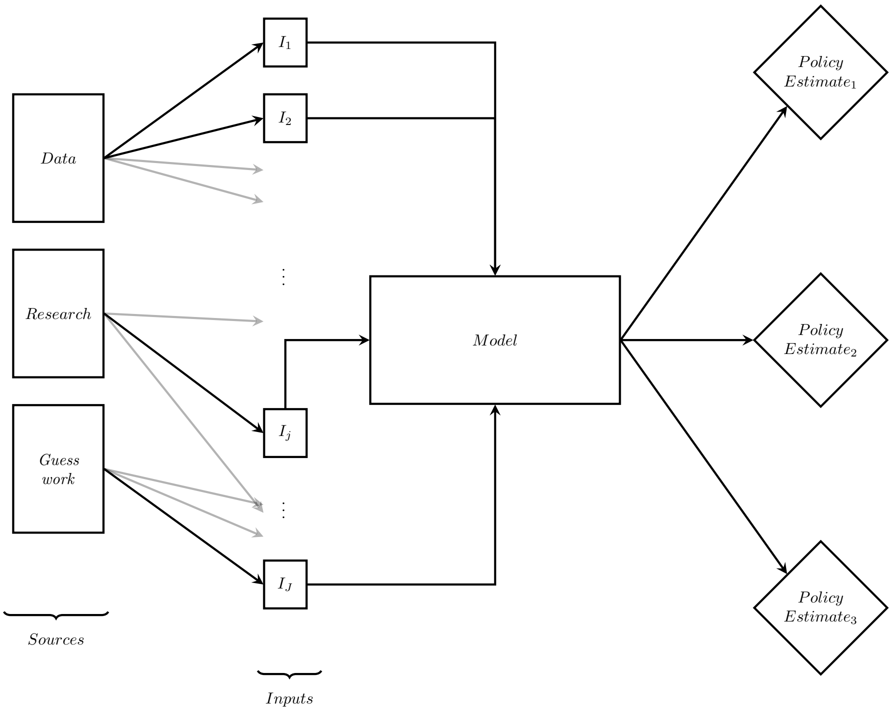


Figure 2: The Process of Policy Analysis

The TR approach to policy analysis develop here consist of three steps: (i) translate guidelines from science to the policy analysis setting, (ii) use the guidelines and state of the art tools (git and dynamic documents) to increase TR in policy analysis (demonstrate with case study), (iii) conduct sensitivity analysis to identify components of the policy analysis where additional knowledge is most policy relevant.

3.1 Adapt TOP Guidelines to Policy Analysis

The guidelines drafted here follow the framework developed in the TOP guidelines for scientific TR (Nosek et al, 2015). The TOP guidelines have eight different dimensions or standards to assess transparency and openness (Citations, Data Transparency, Analytic Methods -code- Transparency, Research Materials Transparency, Design and Analysis Transparency, Preregistration of Studies, Preregistration of Analysis Plans, Replication). Each standard is scored in a scale of four levels, from the lowest score at level 0 and the highest score at level 3.

The goal of the adapted guidelines is to identify standards and levels that would address the problems described in the introduction. For this purpose figure 2 presents all the components to be identified in the process of policy analysis.

The policy analysis process quantifies of costs, benefits and/or distributional effects of a specific policy, all the output of this process will be referred as as *policy estimates*. Policy estimates are produced by a model, good examples of models range in complexity from back-of-the envelope calculation (GiveWell, 2015), micro-simulation models (US Congressional Budget Office, 2014; Agency for Healthcare Research and Quality, 2015), to dynamic choice models (Rothstein, 2015). The type inputs needed for any modeling technique are combinations of research estimates, data and guessed parameters.

3.1.1 Three standards of TR in Policy Analysis

3.1.1.1 Work-flow

Figure 2 represents a stylized version of the work-flow of policy analysis. Inputs are build from three primary sources: data, research, and guess work. The model represents a set procedures applied to inputs in order to quantify the potential effect of the analyzed policy. The model should be characterized by a set of equations and a narrative describing how to apply them. The final policy estimates are the output produced by the model. High TR in the work-flow dimension means that all the components are clearly labeled, together with a clear explanation of how the different pieces feed in to each other.

The goal of any policy analysis is to generate policy estimates, understood as a set of quantities that reflect the best available information regarding objective/positive losses and gains associated with the implementation of a policy. These policy estimates will serve as an input to be valued normatively by policy makers. The policy report should highlight where and why different quantities cannot be compared without a normative assessment. In order to facilitate the normative comparison, all quantities should be reported in the same units.

- **Level 1:** Identify clearly all the policy estimates to be used by policy makers. As different policy makers will weight policy estimates differently (some may focus on the costs, while other on the benefits) it is strongly recommended that all the quantities to be presented in the same unites (e.g. average increase in per-capita income across quantiles of the income distribution).

- **Level 2:** Level 1 + Identify clearly all the inputs to be used in the model and classify their origin as from data, research or guesswork.

- **Level 3:** Level 2 + The complete work-flow should fit a diagram similar to figure 2. Users of the analysis should be able to change specific components with minimal effort, and observe how that change affects the policy estimates.

3.1.1.2 Data

Data used for policy analysis has the purpose of informing decisions in the public sphere. Following new standards (Obama, 2009) these resources should be open by default and a detailed rationale should be made whenever access to the data is restricted. Whenever the raw data cannot be access, the policy analysis should provide access to masked or aggregated data. Additionally, detailed instructions on how to go from the raw data to the masked data should be provided, such that other analysts with access to the raw data could reproduce the exact intermediate data.

- **Level 1:** Policy report states explicitly whether all, some components, or none of the data used in the analysis is available. In the case of differential availability of components, it should be clear which items are fully available, masked, aggregate or non-available at all. Clear instructions should be provided to access each of the available components.
- **Level 2:** Policy report is published with the data. The report and data can be accessed in the same place.
- **Level 3:** Policy report is published with embedded code that calls the data to a repository and changes in the data will produce traceable changes in the report.

3.1.1.3 Methods

Methods in policy analysis should be understood as a detailed set of instructions of all the steps taken to produce the policy estimates. These instructions should describe how each data set, research and guesswork component were use to generate inputs and how those inputs where subsequently used in the model that computes the policy estimates. The intended audience of this instructions should be staffers of policy makers, researchers and other analyst. For this purpose the material should be presented at different levels of depth with a clear narrative connecting all the steps involved.

- **Level 1:** Methods should be clearly described at different levels of detail. A reader of the policy report should be able to understand all the components and reproduce qualitatively similar estimates.
- **Level 2:** Level 1 + all the code that reproduces the exact policy estimates should be available and running.
- **Level 3:** Level 2 + the code and narrative should be clearly legible for different audiences (staffers, researchers, other policy analysts). The code should be in the same document that describes each step, and the users should be able to manipulate all the components of the code and trace its effects on the final policy estimates.

3.2 Apply highest TR to CBO report

Two tools, originated in computer science, facilitate the achievement of the highest TR standards in each of the three dimensions describe above: Dynamic Documents, and Git or Distributed Version Control. A brief description of each tool is presented below with an explanation of how can they help to increase TR in each standard.

Table 1: Summary of Guidelines for Transparent and Reproducible Policy Analysis. Adapted from TOP (Nosek et al, 2015)

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

3.2.1 A Dynamic Document for CBO's report

Dynamic documents (DD) are an important tool promoted in the Data Science community to improve reproducibility of research (Xie, 2015). Build upon the practice of *literate programming* (Knuth, 1992), it consist of a reporting technology that combines the narrative, mathematical modeling, and coding components of the analysis in one single document. This allows to put all the components of the analysis in one place, and update them when new information becomes available in what is known as a one-click reproducible work-flow.

This technology would help to implement the standards describe above up to their highest level. The definition of a DD implies that the data has to be available, published with the report and setup in a way that reacts to changes in the report (level 3 of data standards). A complete DD should also accomplish the highest standard in the methods dimension as it is possible to describe the methodology and the code in detail, and run successfully in different machines (dynamic component). Different users should be able to read it and execute it with minimal effort (level 3 methods standard).

The highest TR in the work-flow standard would not follow immediately from the use of DD, but it should facilitate the achievement of it. Dynamic documents, as opposed to static printed documents, can have different layers of depth, such that different audiences can choose their desired level of detail. At the simplest level the reader/user should be able to visualize a work-flow like the one presented in figure 2.

3.2.2 Distributed Version Control (Git)

Git is a version control system that allows multiple users to edit the same file without losing track of any modification. Since its development in 2005, it has become a universally required tool for software developers, and in the last five years has shown increasing popularity in the research community.

The three main reasons to use of git are: (i) It tracks for all the changes done over any file containing code, rendering obsolete the need for multiple version and names for files. (ii) Allows for multiple users to "clone" a version of the official code and modify it. (iii) Then the contributor can request that her modifications be incorporated to the original file ("pulled" back), and everybody following the original can see exactly all the suggested changes.

Github is a popular website that hosts most of the work using git. The Open Science Framework is another platform that uses git and Github, specifically dedicated for researchers.

As describe in the previous section, a DD provides all the elements to potentially achieve the highest standards of TR in policy analysis. Git provides open access, and allows modifications of the DD, realizing all its potential for TR policy analysis.

3.3 Sensitivity Analysis

When a policy analysis has achieve level 3 of TR in each standard, it is possible to conduct a sensitivity analysis for each component used in the policy analysis. For this exercise it is particularly useful to have all the outputs of the policy analysis (policy estimates) under the same units, and clearly identifying the dimensions to be normatively aggregated by policy makers.

For our case study, the original CBO report presented the benefits and costs in different units: wage gains, wage loses, and total income lost for families with incomes in different poverty line segments (less than one PL, between one and three PL, etc). Taking the DD to the highest level of TR in the work-flow standard implies that this benefits and cost have to be translated in to the same units. For this purpose all the policy estimates are expressed in terms of average per capita income gain/loss, across quintiles of income.

With five quintiles and three type of policy estimates, the dimensionality of the analysis becomes too large even when looking a few parameters. As an illustration of how all dimensions could be condensed in a single number I model the different valuations of hypothetical policy makers using additive weights for each policy estimates and weights to account for different redistributional preferences. The result, labeled as the Normative Policy Estimates (NPE), is a single number that combines all the policy estimates and personal valuations of a given policy maker,

Formally the NPE can be defined as the weighted sum of policy estimates for the wage gain (WG_i), wage loses (WL_i) and balance loses (BL_i) across all individuals, where each PE receives a weight $\omega_{WG}, \omega_{WL}, \omega_{BL}$, and the distributional preferences are a function of the quintile of each individual $\omega_i^d(Q_i, \rho)$:

$$NPE(\omega) = \sum_{i \in N} \left(\omega_{WG} WG_i + \omega_{WL} WL_i + \omega_{BL} BL_i \right) \omega_i^d(Q_i, \rho) \tag{1}$$

with:

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

Q_i represent the quintile in the income distribution (1 the lowest and 5 the highest), and ρ parametrizes the preferences towards redistribution ($\rho < 0$ dislikes redistribution, $\rho > 0$ likes redistribution)

As an illustration, figure 4 presents NPEs for different redistribitutional preferences (ρ) assuming $\omega_{WG} = \omega_{WL} = \omega_{BL}$. In this example a policy maker that values redistribution positively, with $\rho = 0.1$ will see a value of \$9.7 billion dollars over increasing this minimum wage. Conversely, a policy maker that dislikes redistribution ($\rho = -0.1$) will see a NPE of -\$5 billion dollars. The sensitivity analysis is performed over this two sample positions.

4 Results

4.1 CBO Report: Before Increasing TR

Using the guidelines it is possible to assess the current TR of the case study. CBO's report reaches level one across all three standards. It is important to mention that most policy analyses used to support laws or regulations would typically scored at level 0, as the only information available are the policy estimates alone (where all the analysis is presented in statements of the form "it is estimated that policy X will save \$Y millions of dollars"). In this context CBO's report is in the frontier of TR policy analysis.

However, level one of TR still leaves many elements of the analysis unexplained and it requires a large effort to reproduce qualitatively similar results.

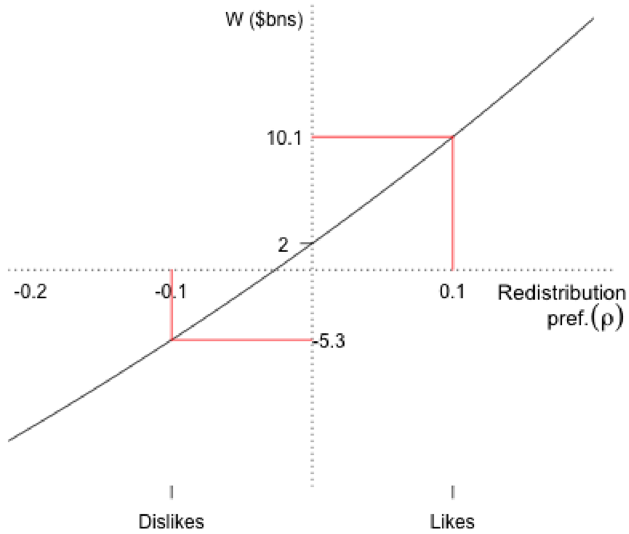


Figure 3: Normative Policy Estimates for different distributional preferences (ρ)

For example, a heavily debated issue after the publication of CBO's analysis was about the effects on employment: 500 thousand jobs loss. As discussed in section 2.2 most of the academic debate can be grouped in to scholars that support a elasticity of labor demand of -0.1 (large effects on employment) and those who support estimates closer to -0.01 (small effects). A quick read of the report would suggest that the former estimates where the ones finally chosen (US Congressional Budget Office, 2014, page 25), but using this elasticity would produce a effect on employment on the order of 300 thousand jobs lost[5]. A more detailed read of the report shows that this parameters was only applied to teenagers, while the elasticity for adults was adjusted by one third (page 28). Incorporating this into the analysis would render a job loss closer to 100 thousand. A detailed review of the report (pages 26-28) would reveal an adjustment that follows (Neumark and Wascher, 2008; Brown, 1999) and would increase the elasticity of teenagers and adults by factors of 3.2 and 19.5 respectively, rendering 1.1 million jobs lost. Only after an exhaustive review of the report it would be clear that the factor used for the final adjustment was 4.5 for both populations (page 28), which renders the reported policy estimate of 500 thousand job lost.

The overall replication process described in the last paragraph required several days of dedicated work, and the final effect on employment depended, in addition to the much debated elasticity, on other components that were largely ignored. A review of the technical discussion following the publication of CBO's report did not reveal any of the elements discussed above.

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

4.2 Reproduce Report With Highest TR

A dynamic document (DD) was produced to increase the TR of CBO's analysis to the highest level and its publicly available online (`http://rpubs.com/fhoces/dd_cbo_test1`)[6]. This section describes how the DD, reaches level three of the guidelines for each standard.

4.2.1 Data

The original version of the report already had achieved level one of TR in the Data standard. The additions made in the DD aimed at putting the data and report together, and allow for automatic update of the data sources.

For example the call of the first data set (Current Population Survey - Outgoing Rotation Group 2013) now happens in section 2.1 of the DD where, for the first time user, the data is downloaded from the web and it requires a one line change of code to repeat the same exercise with different version of the CPS ORG (different year or data repository).

The same methodology was applied for all four data sets used in the analysis (CPS ORG, CPS ASEC, State level minimum wage data base, and 10 year macroeconomic forecast from CBO). If an analyst would like to perform the same analysis but over a different time periods, it would require for each data set to modify the year parameters in the data calls.

4.2.2 Methods

In the methodological description CBO leaves a few unexplained components that where either ignored or guessed, but are explicitly mentioned in the DD. Another dimension for improvement in the Methods standard is the structure of the report. Some analysis is quantitaitive while another fits a more narrative description, and sometimes is not clear which is been described. Here the benefit of hindsight allows to focus the DD on the key methodological components of the analysis that where discussed after the publication of the original report (focusing for example on only one wage raise option instead of two).

The DD also combines the methodological explanations with the code that apply those methods in each step (level 2). And allows for the user/reader to see how the result react to changes in the parameters used (level 3).

4.2.3 Work-flow

Following the structure of figure 2, the DD connects all the sources (data, research and guess work) with the inputs needed to do the analysis. Tables 3 and 4 in the Appendix B list all the components of the policy analysis for the case study and how they connect to each other.

The ability to easily trace the source of each policy estimate makes for one of the strongest contrast between the original report and the DD. Repeating the example of reproducing the effects on employment it is only necessary to find the policy estimate (478 thousand jobs lost) and in the same section the equation behind that calculation is presented, together with a table that contains all the elements needed for such calculation.

[6] Screenshots are provided in the appendix of this paper

An important difference between the original report and the DD, in the work-flow standard, has to do with how the policy estimates are presented to inform the decision of policy makers. To achieve the highest TR in the work-flow standard, all the information needed by the policy makers should be contain in a single visualization or table. This output should identify all the positive components produced by the analysis to be weighted differently by different policy makers. But among analysts a consensus should be achieved on how to present such outcome, and it format should be invariant to future versions of the report. The current version of the DD presents all the policy estimates in the same unit: average gain/loss per-capita, and displays the distributional effects, in quintiles, of each policy estimate. This output is presented in figure 4.

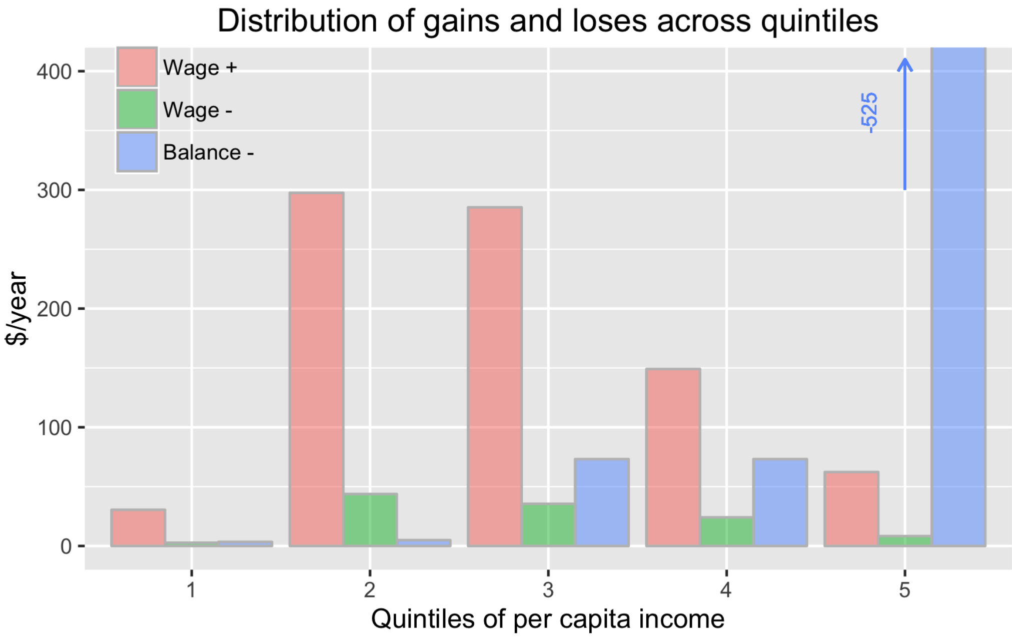


Figure 4: Distributional display of policy estimates

4.3 Sensitivity analysis

One of the main purposes of the DD is that, after achieving the highest levels of TR, an arbitrary large number of sensitivity analysis can be performed to assess how the components of the analysis affects the final result. Here I perform a number of sensitivity analyses, but the goal is for the DD to allow other users, with minimal effort, to perform different sensitivity analysis of their own interest.

For illustration purposes, first the variation of a couple of inputs is assess over the final set of policy estimates that are used to inform the policy makers (15 quantities: 3 policy estimates across 5 quintiles). Then the example of hypothetical preferences for different policy makers is used to perform the sensitivity analysis over a larger number of inputs and applying the same variation to every input.

Given that most of the academic debate on minimum wage has been between the schools of large effects on employment (with an elasticity for teenagers of -0.1) and small effects on employment (elasticity of -0.01), a natural candidate for first sensitivity analysis is the elasticity of labor demand for teenagers. The results are presented in figure 5. To facilitate the comparison with figure 4 the margins of the vertical axis have held the same.

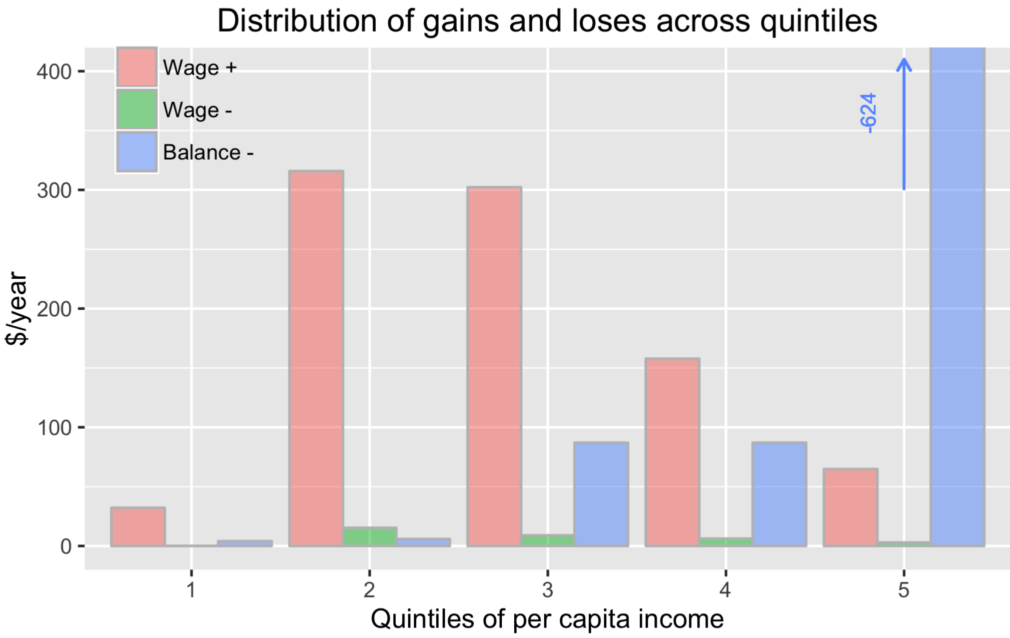


Figure 5: Change in Elasticity of Labor Demand: from $\eta_{lit}^{teens} = -0.1$ to $\eta_{lit}^{teens} = -0.01 (\Delta^{-} 90\%)$

The variation in the set of policy estimates seems small if all gains and loses are weight equally. This suggests that the academic discussion between an elasticity of -0.1 or -0.01 might not be as policy relevant, or that a particularly large normative weight is given to the loses of the group of workers that lose their jobs.

For a comparison figure 6 presents the output when the distribution of balance loses is changed from[7] $(1\%, 29\%, 70\%)$, to $(20\%, 40\%, 40\%)$. In this case there is a substantive change in how some of the loses are distributed and seems more likely that differences between figure 4 and 6 would have a larger effect on the decision of policy makers more than differences between 4 and 5.

Next, the sensitivity analysis is carried out over a larger set of the inputs identified in the DD (table 3 in Appendix B) and its effect is evaluated over the hypothetical preferences of two type of policy makers: one that favors redistribution and one that does not ($NPE(\rho = 0.1)$ and $NPE(\rho = -0.1)$ of equation 1 respectively).

[7]1% the total wage gain is paid by individuals with income below the poverty line, 29% by those with income between one and 6 poverty lines, and 70% by those with incomes greater than 6 poverty lines

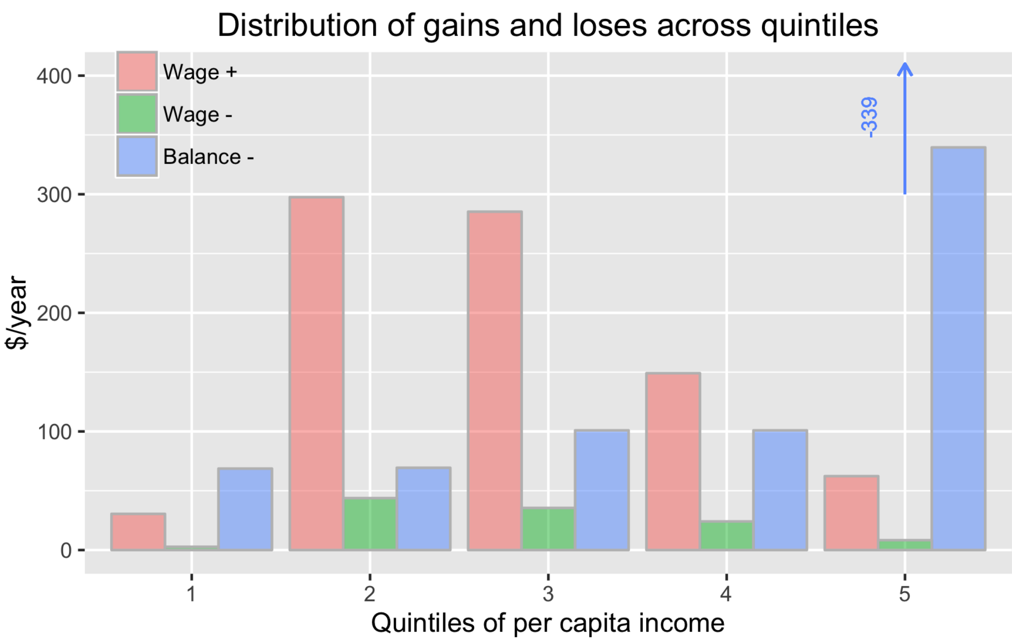


Figure 6: Change in Distribution of Balance Loses: from $(1PL, 6PL) \sim (1\%, 29\%, 70\%)$ to $(20\%, 40\%, 40\%)$

For each parameter two type of perturbations were applied: a 10% increase and a 10% decrease from its current level for each parameter. This methodology was applied to all the parameters in the table with the exception than those that describe the distribution of balance loses. In the sensitivity analysis, the parameters for the distribution of balances loses were chosen to reflect three scenarios: a higher share of the losses paid by the wealthy $(1\%, 4\%, 95\%)$, a slightly less "progressive" loss scheme $(5\%, 35\%, 60\%)$, and a flat distribution of loses $(40\%, 54\%, 6\%)$, and a uniform distribution of balance losses.

Table 2 suggests that, in addition to the elasticity of labor demand, there are many other components of the policy analysis that can have consequential effect on the final decision made by policy makers.

For example, changes in the current values of the ripple effects inputs would play a pivotal role in either of the two hypothetical positions. A 10% variation in the scope parameters (the range of the ripples), from $(8.7, 11.5)$ to $(7.8, 12.7)$ would increase the policy makers' valuations in 37% for those against, and in 21% for those in favor of raising the minimum wage.

It is important to acknowledge that the academic debate around the elasticity of labor demand would represent a 90% reduction on the current value used in the analysis. This implies changes in valuations around 36% and 18%, making it a consequential debate for policy purposes. What the results in table 2 suggest is that other components play an equally important role, and yet much less is know about them.

The ripple effects could have a much larger or narrower scope and intensity, as the literature around it is scarce. All the parameters in the Guess Work

Table 2: $\%\Delta W$ 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%
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%	

category can have a much wider range as almost nothing is known about them.

The most consequential parameters by far are the ones that describe the distribution of balance loses. Almost nothing is known about who finally carries the burden of a raise in the minimum wage. Understanding such distribution should be a priority in a evidence-based policy debate around the minimum wage.

The sensitivity analysis performed here represents a small fraction of all the variation that can be studied once, particularly if combinations of variations are considered. The goal is to motivate readers to clone their own version of the DD and perform multiple analysis of interest. This is one of the main benefits of the open-source feature of the dynamic document develop for this paper.

5 Discussion

This paper translates and implements guidelines and tools for transparency and reproducibility (TR) in science, into policy analysis. The goal of this methodological innovation was to propose an additional solution to the critique of low credibility of policy analysis (Manski, 2013). Increasing TR in policy analysis increases its credibility, helps to provide a clear connection between research and policy analysis, allows for systematic improvement and automation in specific and recurrent policy analyses. Drawing a parallel between the reproducibility/credibility crisis in science, made it possible to identify similar solutions. After translating standards of TR in to the policy analysis setting, as a case study I implemented the highest level of TR in to a report from the Congressional Budget Office on the effects of raising the minimum wage.

The results from the case study show how a dynamic document in open source format, the highest standard of TR, can draw a clear connection between research inputs and the output of a policy analysis. The transparency component helped identify potential weaknesses in the original policy analysis, while the reproducibility component allowed for a comprehensive sensitivity analysis that shed light on where new knowledge could be the most valuable. As an example, the sensitivity analysis suggests that, from perspective of policy relevancy, the research agenda on the effects of minimum wage on employment is over-studied relative to other areas like the distribution of loses used to pay for the increase wages.

The open source nature of the dynamic document aims to provide the foundations for a constructive debate around the technical issues of a recurrent policy issue. It allows for updates in short term and long term dimensions. The vast majority of a policy analysis is to "put down fires", i.e. address pressing issues that need some type of immediate solution. For this reason the technical contributions in the short term category should accept the proposed model as the correct one and oversee its correct implementation. In parallel, contributors to the DD can propose structural modifications to the analysis, to be address in the next cycle of the policy debate.

If this approach to policy analysis were to become the status quo, benefits would be observe in three dimensions. First the cost of producing the next (or marginal) report would be reduced substantially. Second, any modification to the original report would be incremental, as oppose to arbitrary, as the policy analysis debate discussion would now have a framework to systematically improve upon previous work. Third, as clarity is added the positive elements of a policy discussion, it would be easier to have normative policy debate that clarifies the positions of different policy makers.

Finally a key issue that is not addressed in this papers regards to who should be responsible for the implementation of this approach. Policy analysts face strong time and resource constrains, and adding a set of protocols and techniques seems not plausible. Those best suited of this task should have a less stringent resource constrain, but be closely related to the the policy issue and analysis. Possible candidates are banks of knowledge as proposed by Clemens and Kremer (2016), public policy schools, think tanks or expert commissions.

References

Abbasi K (2014) The missing data that cost $20 bn. BmJ 348:g2695

Agency for Healthcare Research and Quality (2015) Evaluating the benefits and harms of colorectal cancer screening strategies: A collaborative modeling approach

Brown C (1999) Minimum wages, employment, and the distribution of income. Handbook of labor economics 3:2101–2163

Card D, Krueger AB (2015) Myth and measurement. Princeton University Pres

Christensen G, Soderberg C (2015) Manual of best practices in transparent social science research. Berkeley, CA: University of California Retrieved April 23:2016

Clemens J (2015) The minimum wage and the great recession: Evidence from the current population survey. Tech. rep., National Bureau of Economic Research

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

Collaboration OS, et al (2015) Estimating the reproducibility of psychological science. Science 349(6251):aac4716

Dube A, Lester TW, Reich M (2010) Minimum wage effects across state borders: Estimates using contiguous counties. The review of economics and statistics 92(4):945–964

Franco A, Malhotra N, Simonovits G (2014) Publication bias in the social sciences: Unlocking the file drawer. Science 345(6203):1502–1505

GiveWell (2015) Givewell's cost-effectiveness analyses — givewell. `http://www.givewell.org/international/technical/criteria/cost-effectiveness/cost-effectiveness-models`, (Accessed on 04/14/2016)

Gray J (2014) Towards a genealogy of open data. In: The paper was given at the General Conference of the European Consortium for Political Research in Glasgow

Ioannidis JP (2005) Why most published research findings are false. PLoS Med 2(8):e124

Jager LR, Leek JT (2013) An estimate of the science-wise false discovery rate and application to the top medical literature. Biostatistics p kxt007

Knuth DE (1992) Literate programming. CSLI Lecture Notes, Stanford, CA: Center for the Study of Language and Information (CSLI), 1992 1

Leamer EE (1983) Let's take the con out of econometrics. The American Economic Review 73(1):31–43

Manski CF (2013) Public policy in an uncertain world: analysis and decisions. Harvard University Press

Merton RK (1973) The sociology of science: Theoretical and empirical investigations. University of Chicago press

Miguel E, Camerer C, Casey K, Cohen J, Esterling KM, Gerber A, Glennerster R, Green DP, Humphreys M, Imbens G, et al (2014) Promoting transparency in social science research. Science 343(6166):30–31

Neumark D, Wascher WL (2008) Minimum wages. MIT Press

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

Nutley SM, Walter I, Davies HT (2007) Using evidence: How research can inform public services. Policy press

Obama B (2009) Transparency and open government. Memorandum for the heads of executive departments and agencies

Peng RD (2011) Reproducible research in computational science. Science (New York, Ny) 334(6060):1226

Rosenthal R (1979) The file drawer problem and tolerance for null results. Psychological bulletin 86(3):638

Rothstein J (2015) Teacher quality policy when supply matters. American Economic Review 105(1):100–130

Simonsohn U, Nelson LD, Simmons JP (2014) P-curve: A key to the file-drawer. Journal of Experimental Psychology: General 143(2):534

Smith V, Prescott E, Fama E, 497+ (2014) Economist letter against raising the minimum wage. http://tinyurl.com/kx3xphr, (Accessed on 10/31/2016)

Snilstveit B, Vojtkova M, Bhavsar A, Gaarder M (2013) Evidence gap maps–a tool for promoting evidence-informed policy and prioritizing future research. World Bank Policy Research Working Paper (6725)

US Congressional Budget Office (2014) The effects of a minimum-wage increase on employment and family income

Vivalt E (2015) How much can we generalize from impact evaluations. Unpublished manuscript

White House (2013) Executive order–making open and machine readable the new default for government information. White paper

White House (2014) Congressional budget office report finds minimum wage lifts wages for 16.5 million workers — whitehouse.gov. http://tinyurl.com/kkd8fcu, (Accessed on 10/31/2016)

Xie Y (2015) Dynamic Documents with R and knitr, vol 29. CRC Press

6 Appendices

6.1 Appendix A: Dynamic Document of Case Study

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.

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

2 Employment effects
2.1 Data, wages, and forecast
2.2 Get the N
2.3 Get the $\eta \times \Delta w$
2.4 Other factors
2.5 Computing effects on employment
3 Distributional effects
3.1 Computing Family income
3.2 Imputing policy effects
3.3 Computing family income under status quo and minimum wage increase
3.4 Other considerations
4 Results

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")
```

2.5 Computing effects on employment

Putting all elements together we get:

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

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
$\widetilde{\eta_{w\leq MW}}$	-0.15	-0.45

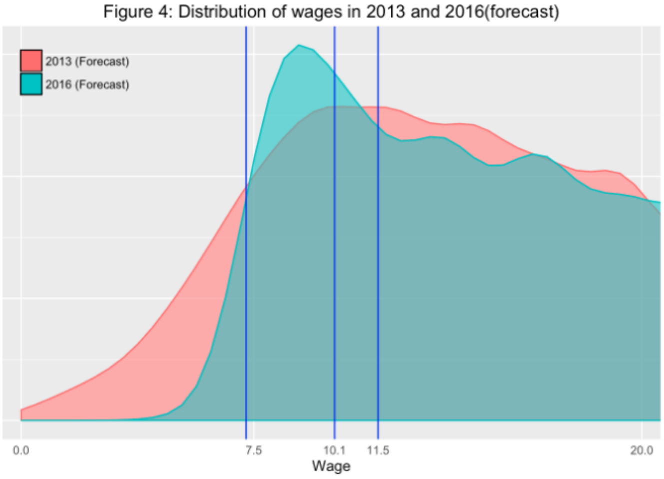
Using all the components described above we get $\widehat{\Delta^- E} =$ -478 thousand jobs. The report however computes F^g_{adjs} in a different fashion and gets a value of 4.5 (when computing the values of F^g_{adjs} from the table below - as oppose to using historical values - we get $\widehat{\Delta^-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

3 Distributional effects

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

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

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

6.2 Appendix B: Detail Link Between Sources-Inputs-Model-Policy Estimates for Case Study

Table 3: Sources and inputs connection

Source	Input
Data	
CPS ORG 2013 (CEPR version)	Number of salary workers in 2013; Fraction of workers below the new minimum wage; Average wage variation for those below the new min wage
CPS ASEC 2012 (CEPR version)	Wages; Non-Wage Income; Household size; Hours/weeks worked
State level Min. Wage (DOL)	Trends in state min. wage
10-year economic forecast (CBO)	Number of workers predicted by 2016 (in 2013); Wage growth in by 2016 (in 2013)
Research	
Elasticity of labor demand for teenagers	-0.1
Ripple effects	From \$8.7 to \$11.5 with a 50% "ripple"
Non compliance rate	
Guess Work	
Extrapolation factor from teenagers to adults	1/3
Net benefits	\$2 billion
Adjustment to account for effective wage variation and affected population	4.5
Aggregate consumption effects on employment	40,000 new jobs
Distribution of balance loses	1% if income $\leq 1PL$, 29% if income $\in (1PL, 6PL)$, 70% if income $\geq 6PL$
Fraction of wage loses used to pay for wage gains	1
Job killing process: fraction of jobs	Cut wages in half for twice the number of jobs destroyed

The connection between inputs and the methods is described in detail in the DD. All inputs listed in table **??** can be found in each equation characterizing the methodology.

Table 4: Sources and inputs connection

Model	Policy estimate (per quintile)
Predicted household income with and without min wage increase. **Depends on**: $\widehat{N^g_{final}}, P_{\hat{w} \leq MW^1 \| g}, \overline{\%\Delta w^g}, \alpha^g_1$, $dF_w, dF_{nw}, N_h, \hat{w}, \hat{h}, MW^s_t, \hat{g_N}, \hat{g_w}, \hat{g_{nw}}$, $\eta^{lit}_{teen}, R_{lb}, R_{ub}, R_I, F_{ex}, F_{adj}, \hat{OF}$	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**: $\widehat{N^g_{final}}, P_{\hat{w} \leq MW^1 \| g}, \overline{\%\Delta w^g}, \alpha^g_1$, $dF_w, dF_{nw}, N_h, \hat{w}, \hat{h}, MW^s_t, \hat{g_N}, \hat{g_w}, \hat{g_{nw}}$, $\eta^{lit}_{teen}, F_{ex}, F_{adj}, \hat{OF}$	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. ($\overline{BL_q}$)

6.3 Appendix C: Additional Plots for Sensitivity Analysis

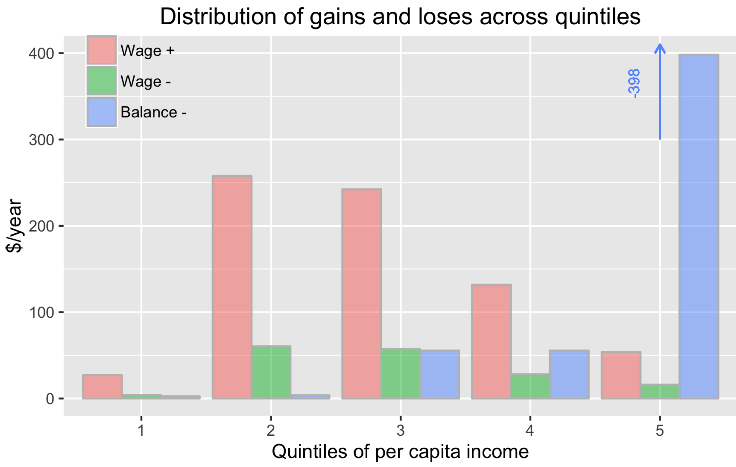


Figure 7: Change in No Compliance Rate: from $\alpha_1 \approx 15\%$ to $\alpha_1 \approx 23\%(\Delta^+ 50\%)$

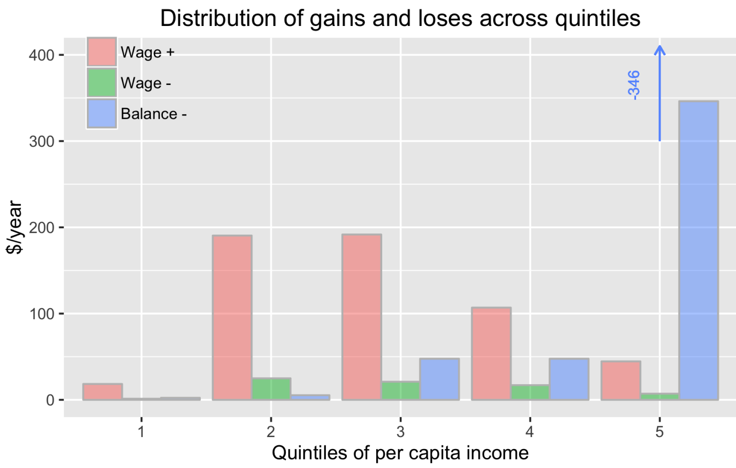


Figure 8: Change in States Minimum Wage: from $\overline{MW_s} \approx 8$ to $\overline{MW_s} \approx 9.6(\Delta^+ 20\%)$