The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$115	1-year subscription	$145
2-year subscription	$210	2-year subscription	$270
3-year subscription	$285	3-year subscription	$375
1-year student subscription	$ 85	1-year student subscription	$115
1-year institutional subscription	$345	1-year institutional subscription	$375
2-year institutional subscription	$625	2-year institutional subscription	$685
3-year institutional subscription	$875	3-year institutional subscription	$965
Electronic only		**Electronic only**	
1-year subscription	$ 85	1-year subscription	$ 85
2-year subscription	$155	2-year subscription	$155
3-year subscription	$215	3-year subscription	$215
1-year student subscription	$ 55	1-year student subscription	$ 55

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.


The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, STATA, Stata Press, Mata, MATA, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2015)
15, Number 2, pp. 397–410

Person-centered treatment (PeT) effects: Individualized treatment effects using instrumental variables

Anirban Basu
Departments of Pharmacy, Health Services, and Economics
University of Washington
Seattle, WA
and National Bureau of Economic Research
Cambridge, MA
basua@uw.edu

Abstract. I describe a command, `petiv`, that uses a local instrumental-variables (LIV) approach to estimate person-centered treatment effects for a variety of specifications for the LIV estimand as outlined in Basu (2014, *Journal of Applied Econometrics* 29: 671–691). The `petiv` command creates a new variable in the dataset that contains the person-centered treatment effects for each individual in the dataset. However, the command takes the validity of the instrumental variables and the specification of the LIV estimand as given. Appropriateness of these features of an LIV analysis should be determined before running the `petiv` command. The individual effects can be used to answer distributional questions and can also be easily aggregated to obtain mean treatment-effects estimates.

Keywords: st0385, petiv, local instrumental-variables methods, person-centered treatment effects, treatment-effects heterogeneity

1 Introduction

In the evaluation literature, nuanced treatment effects are popularly characterized by conditional average treatment effects (CATEs), where an average treatment effect (ATE) is estimated conditional on certain values of observed covariates over which treatment effects vary. For example, if age is the only observed risk factor, one can establish a conditional effect of surgery versus active surveillance on mortality for patients of age 60 years diagnosed with clinically localized prostate cancer. This is an average effect for all 60-year-olds in this condition. However, does this estimate apply to all men with clinically localized prostate cancer at age 60 years? Certainly not, because there may be many other factors that determine heterogeneity in treatment effects in this population. For example, clinical stage and grade of cancer not only determine overall survival but also may determine differential effects from alternative treatments. To the extent that all potential moderators of treatments effects are observed in the data, a nuanced CATE can be established conditioning on values of all of these factors.

In most practical analyses, all moderators of treatment effects are not observed. Many of these moderators are yet to be discovered and, hence, remain unknown to scientific knowledge. They are typically represented by the pure stochastic error term in statistical data analysis. However, there are some moderators within the purview of scientific knowledge that remain unmeasured in the data at hand. For example, most randomized studies, which rely on randomization to equate the distribution of all of these factors across the randomization arms, forgo measurement of several factors in the interest of time and expense.

In observational studies, these unmeasured moderators of treatment effects not only cause selection bias (Newhouse and McClellan 1998) but also play a vital role in generating essential heterogeneity because they are often observed by individuals and acted upon while making treatment selection (Heckman and Vytlacil 1999). An entire genre of methods, including methods based on local instrumental-variables (LIV) approaches, has been developed to estimate policy-relevant and structurally stable mean treatment-effect parameters in the presence of essential heterogeneity (Heckman and Vytlacil 2001, 2005). Basu and colleagues introduced these methods to the health economics literature, where essential heterogeneity is widespread and instrumental-variables (IV) methods are gaining popularity (Basu et al. 2007; Basu 2011).

LIV methods can seamlessly explore treatment-effects heterogeneity across both observable characteristics and unobserved confounders and can be used to establish CATEs based on observed factors. Basu extended this literature by developing a new individualized treatment-effects concept called person-centered treatment (PeT) effects, which can also be estimated using LIV methods (Basu 2014). This new treatment-effects concept is more personalized than CATEs because it takes into account individual treatment choices and the circumstances under which people make those choices in an observational data setting to predict their individualized treatment effects. There are several intuitive aspects about the PeT effects:

1. They help one comprehend individual-level treatment-effects heterogeneity better than CATEs and can explain a larger fraction of the individual-level variability in treatment effects than CATEs.

2. They are better indicators for the degree of self-selection than CATEs. They are better predictors of true treatment effects at the individual level for both the positive predictive value and the negative predictive value (Basu 2014).

3. All mean treatment-effect parameters, such as the ATE and the effect on the treated, can be easily computed from PeT effects.

2 PeT effects

2.1 Intuitive ideas behind essential heterogeneity, marginal treatment effects, and PeT effects

To provide the intuition behind these concepts, we start with a stylized example. It is widely known that the effectiveness of surgery is heterogeneous across men with prostate cancer. Let's assume that surgery is more effective for younger men and for men with high baseline prostate-specific antigen (PSA) values. Moreover, let's assume choice of treatment (surgery or watchful waiting) is influenced by age and baseline PSA of the patient.

However, if one obtains a sample of men diagnosed with prostate cancer from Surveillance, Epidemiology, and End Results (SEER)-Medicare, "PSA values" will not be measured. Suppose the available dataset contains the following information for each patient: age, treatment, whether the patient was cured, and a characteristic that all agree is a powerful and valid instrument. Let the instrument be a continuous variable, for example, distance to hospital. The greater the distance to the hospital, the less likely the use of surgery. Distance is statistically independent of all risk factors that determine outcomes, a condition that is met when the two underlying principles are met; that is, distance does not directly affect outcomes, and individuals do not select residences such that their distances from the hospital are correlated with their risk factors. Note that one or more continuous instruments are needed to identify PeT effects. Without continuous instruments, only treatment-effects bounds (Shaikh and Vytlacil 2011) and the average effects on some compliers (Frandsen, Frölich, and Melly 2012; Frölich and Melly 2013) can be estimated.

Here a traditional naïve regression would produce biased estimates of the ATE and also of the CATEs for old and young groups because of the endogeneity of the treatment status caused by the missingness of PSA values in the analysis.

A traditional IV analysis, using a strong and valid IV, will also produce biased estimates for ATE and CATEs because of essential heterogeneity, which suggests that the treatment effects vary over unobserved confounders (in this case, the PSA values).

A LIV approach can be used to overcome these issues when a continuous instrument is available. LIV methods are used to estimate the marginal treatment-effects (MTEs) parameters. MTEs are the effects for individuals for whom the influence of the observed characteristics (old age and distance to hospital) balance with the influence of the unobserved confounders (PSA level) on treatment choice such that they are indifferent to choosing between using surgery or watchful waiting. To estimate an MTE, LIV methods compare the outcomes of two groups of, say, young patients, where one group is staying at a distance d from the hospital and the other at a distance $d+\varepsilon$, ε representing an epsilon (very small) change in distance. These two groups of patients should have identical distributions of risk factors (observed and unobserved). If distance is a valid IV, then, by definition, it is independent of all risk factors affecting outcomes. Treatment choices are affected differently between these two groups only through the costs of traveling

the extra distance. Therefore, any difference between the average outcomes between these two groups must be a result of there being some people in these groups (who are identical in their risk-factor distributions) with different treatment choices caused by differences in distance. However, because the difference in distance is very small, this difference in outcomes can be attributed to the effect of treatment on a margin of patients who were indifferent between two treatment options but made different treatment choices based on the small perturbation of the IV (that is, distance). For this margin of patients, we can quantify a normalized level of unobserved confounders because unobserved confounders must balance the observed levels for the patients to be indifferent between treatment choices.

Here normalized means a scalar score that represents a balancing score for unobserved risk factors, irrespective of their empirical distributions. Technically, (2) below shows that such a scalar score is always distributed uniform $(0, 1)$. This scalar score is conceptualized by constructing a scalar propensity score with many observed risk factors with varying empirical distributions.

Similarly, for another dyad of distances, d' and $d' + \varepsilon$, one can estimate another MTE, which reflects the causal treatment effect of patients at another level of unobserved confounders. In this way, a full schedule of MTEs can be estimated that vary over the unobserved-confounders levels (that is, PSA values) given any level of the observed confounder (that is, age). Intuitively, LIV methods estimate these MTEs by first estimating a control function, which models how the observed outcome varies over the observed risk factors, the IV-dependent estimated propensity to choose treatment, the interactions between them, and nonlinear polynomials of the propensity score. The partial derivate of the outcome, as characterized by the control function, with respect to the IV-dependent propensity score (reflecting epsilon changes) estimates the MTE evaluated at specific values of the scalar unobserved risk-factor levels.

Once MTEs are estimated over the observed and unobserved levels, they can then be easily aggregated to form meaningful treatment-effect parameters such as the ATE, CATEs, treatment on treated (TT), and treatment on the untreated (TUT); they can also be used to study heterogeneity in effects by using PeT effects. The PeT effect for a patient in this stylized sample is conditioned on that individual's age, and his or her age-specific MTEs are also averaged over a distribution of PSA levels that conforms with the individual's observed choice of surgery or watchful waiting. Thus these are deemed to be the personalized effects for this patient.

2.2 Formal models behind essential heterogeneity, MTEs, and PeT effects

We restrict our discussion to two treatment states—the treated state denoted by $j = 1$ and the untreated state denoted by $j = 0$—because `petiv` is designed only for binary treatments. However, theoretical extensions to multiple ordered treatments are possible (Heckman, Urzua, and Vytlacil 2006). The corresponding potential individual outcomes in these two states are denoted by Y_1 and Y_0. We assume

$$Y_1 = \mu_1(\mathbf{X}_O, \mathbf{X}_U, \vartheta) \quad \text{and} \quad Y_0 = \mu_0(\mathbf{X}_O, \mathbf{X}_U, \vartheta) \tag{1}$$

where $\mathbf{X}_O$ is a vector of observed random variables; $\mathbf{X}_U$ is a vector of unobserved random variables, which are also believed to influence treatment selection (they are the unobserved confounders); and ϑ is an unobserved random variable that captures all remaining unobserved random variables. $(\mathbf{X}_O, \mathbf{X}_U) \amalg \vartheta$ and $\mathbf{X}_O \amalg \mathbf{X}_U$, where $\amalg$ denotes statistical independence.

We assume individuals choose to be in state 1 or 0 (prior to realizing the outcome of interest) according to

$$D = 1 \quad \text{if} \quad \mu_D(\mathbf{X}_O, \mathbf{Z}) - U_D > 0 \tag{2}$$

where $\mathbf{Z}$ is a (nondegenerate) vector of observed random variables (instruments) influencing the decision equation but not the potential outcome equations, μ_D is an unknown function of $\mathbf{X}_O$ and $\mathbf{Z}$, and U_D is a random variable that captures $\mathbf{X}_U$ and all remaining unobserved random variables influencing choice. By definition, $U_D \amalg \vartheta$, which also defines the distinction between $\mathbf{X}_U$ and ϑ in (1). Equations (1) and (1) represent the nonparametric models that conform to Imbens and Angrist's (1994) independence and monotonicity assumptions needed to interpret IV estimates in a model of heterogeneous returns (Vytlacil 2002). As in Heckman and Vytlacil (1999), we can rewrite (1) as

$$D = 1 \quad \text{if} \quad P(\mathbf{X}_O = \mathbf{x}_O, \mathbf{Z} = \mathbf{z}) > V \tag{3}$$

where $V = F_{U_D}(U_D | \mathbf{X}_O = \mathbf{x}_O, \mathbf{Z} = \mathbf{z}), P(\mathbf{x}_O, \mathbf{z}) = F_{U_D|\mathbf{X}_O,\mathbf{Z}}\{\mu_D(\mathbf{x}_O, \mathbf{z})\}$, and F represents a cumulative distribution function. Therefore, for any arbitrary distribution of U_D conditional on $\mathbf{X}_O$ and $\mathbf{Z}$, by definition, $V \sim \text{Unif}[0, 1]$ conditional on $\mathbf{X}_O$ and $\mathbf{Z}$. Under regular IV assumptions, MTEs can be identified by

$$\frac{\partial E_\vartheta(Y | \mathbf{X}_O = \mathbf{x}_O, \mathbf{Z} = \mathbf{z})}{\partial p} = E_\vartheta\{(Y_1 - Y_0) | \mathbf{X}_O, V = v\} = \text{MTE}(\mathbf{x}_O, v)$$

where $Y = D \times Y_1 + (1 - D) \times Y_0$ is the observed outcome and $v = P(\mathbf{x}_O, \mathbf{z})$.

MTE is perhaps the most nuanced estimable effect. It identifies an effect for an individual who is at the margin of choice such that one's levels of $\mathbf{X}_O$ and $\mathbf{Z}$ are just balanced by one's level of V (which includes $\mathbf{X}_U$); that is, $P(\mathbf{x}_O, \mathbf{z}) = v$. Basu (2014) extends the LIV methods to identify PeT effects, which, for persons who choose treatment, follow

$$\begin{aligned} E_{\mathbf{X}_U|\mathbf{X}_O,P(\mathbf{Z}),D} E_\vartheta\{Y_1 - Y_0 | \mathbf{x}_O, P(\mathbf{z}), D = 1\} &= E\{Y_1 - Y_0 | \mathbf{x}_O, V < P(\mathbf{z})\} \\ &= P(\mathbf{z})^{-1} \int_0^{P(\mathbf{z})} \text{MTE}(\mathbf{x}_O, v) dv \end{aligned}$$

Similarly, the conditional effect for a person who did not choose treatment is obtained by integrating MTEs over values of V greater than $P(\mathbf{z})$.

Conceptually, a PeT effect is also a weighted version of an MTE. This is because an MTE is the treatment effect of a hypothetical individual who is at the margin of choice

because the individual's propensity to choose treatment based on $\mathbf{X}_O$ and $\mathbf{Z}$ is balanced by the propensity to select the alternative based on V. As the value of V is changed from this point, this individual would choose either the treatment or the alternative. The PeT effect for a real individual is then the average of MTEs, with the same $\mathbf{X}_O$ and $\mathbf{Z}$ levels as those for this real individual, over those values of V that correspond to the real individual's own treatment choice. That is, for any given individual, the first step is to identify the specific margins of V where that individual may belong given his or her individual values of $\mathbf{X}_O$, $P(\mathbf{Z})$, and D. Then, the average of MTEs over those margins, but not over all as in CATEs, is calculated to obtain a PeT effect. Further details can be found in Basu (2014).

3 A numerical algorithm to compute PeT effects

To estimate PeT effects of a binary exposure on an outcome, one should perform the following steps:

1. Check the strength and validity of the IVs by using standard methods. Run the first stage by regressing the indicator for exposure (D) against observed factors ($\mathbf{X}$) and the instrument ($\mathbf{Z}$) by using a probit/logit or other model appropriate for a binary outcome. Propensity score $\widehat{p}(\mathbf{x}, \mathbf{z})$ is predicted for every individual in the dataset.

2. Ensure that $\widehat{p}(\mathbf{x}, \mathbf{z})$ has mass at any value (rounded to 0.01) for both levels of exposure. Observations corresponding to particular values of $\widehat{p}(\mathbf{x}, \mathbf{z})$ that do not meet this criteria are dropped.

3. Denote $\min p = \min\{\widehat{p}(\mathbf{x}, \mathbf{z})\}$ and $\max p = \max\{\widehat{p}(\mathbf{x}, \mathbf{z})\}$.

4. Determine the appropriate specification for $g(Y) = \alpha_0 + \alpha_1 \times \mathbf{X} + \alpha_3 \times \widehat{p} \times \mathbf{X} + K(\alpha; \widehat{p})$, where the link function $g(\,)$ and polynomial function of $\widehat{p}(\mathbf{x}, \mathbf{z})$, $K(\,)$, are determined using various goodness-of-fit tests.

5. Specify the LIV estimand through the `petiv` command, as follows:

    a. Run the second-stage LIV estimand for outcome Y by using a user-specified regression model, $g(Y) = \alpha_0 + \alpha_1 \times \mathbf{X} + \alpha_3 \times \widehat{p} \times \mathbf{X} + K(\alpha; \widehat{p})$.

    b. Draw 1,000 deviates $u \sim \text{Uniform}[\min p, \max p]$.

    c. Perform numerical integration. For each individual i, do the following:

        i. Compute $d\widehat{g}(\cdot)/d\widehat{p}$, and evaluate it by replacing $\widehat{p}(\mathbf{x}, \mathbf{z})$ with each value of u so that there are 1,000 values of $d\widehat{g}(\cdot)/d\widehat{p}$ for each individual i.

        ii. Compute $D^* = \Phi^{-1}\{\widehat{p}(\mathbf{x}, \mathbf{z})\} + \Phi^{-1}(1 - u)$, also generating 1,000 values for each individual i, where $\Phi()$ is the cumulative normal distribution function.

iii. Compute PeT effects by averaging $d\widehat{g}(\cdot)/d\widehat{p}$ over values of u for which $(D^* > 0)$ if $D = 1$ or by averaging $d\widehat{g}(\cdot)/d\widehat{p}$ over values of u for which $(D^* \leq 0)$ if $D = 0$.

Estimated PeT effects provide us with individualized treatment effects. Mean treatment-effect parameters can also be computed using these PeT effects. Averaging PeT effects over all observations gives an empirical estimate of the ATE. Averaging PeT effects over $D = 1$ or $D = 0$ provides us with the effect on the treated (TT) and the effect on the untreated (TUT), respectively.

3.1 Inference

Standard errors for individual PeT effects and ATEs can be obtained via bootstrap. Specifically, saving the average PeT effects for each individual (that is, with specific individual ID), where the average is taken if the same individual is sampled more than once within the same bootstrap data replicate, from each bootstrap replicate would help build a distribution of PeT effects for that individual.

4 The petiv command

The `petiv` command performs the LIV estimation and the numerical integration that follows when calculating PeT effects, as previously outlined. At the end of the command, a new variable called `pet_`*depvar* is created in the dataset that stores the individualized effects for each person in the sample.

4.1 Syntax

`petiv` *depvar varlist* [*if*] [*in*], `trt(`*varname*`) ps(`*varname*`)`
`cmd(`*command_name*`)` [`degree(#) controls(`*varlist*`) display` *options*]

`petiv` expects the data to be in the conventional form as in any other regression analysis. It requires specification of a dependent variable and at least one covariate that is not the treatment indicator; that is, it does not fit a constant-only model.

4.2 Options

`trt(`*varname*`)` specifies a binary treatment variable. `trt()` is required.

`ps(`*varname*`)` specifies the variable that contains the estimated propensity score of treatment choice as a function of independent risk factors and IV. `ps()` is required.

`cmd(`*command_name*`)` specifies the regression command to be used to specify the control function that estimates the MTEs and the PeT effects. All interactions of *varlist*

with `ps()` are accounted for. *command_name* is one of `probit`, `logit`, `glm`, `pglm`, or `regress`. `cmd()` is required.

`degree(#)` specifies the degree of polynomial for `ps()` that will be used in the control function. The default is `degree(1)`.

`controls(`*varlist*`)` specifies the list of variables that will be adjusted for in the control function, but no interaction with `ps()` will be used. The default is an empty list. The variables listed within this option should not be listed under *varlist* after *depvar*.

`display` displays the regression results from the estimation of the control function.

options can be any options corresponding to the *command_name*.

5 An empirical example using PeT

This example follows the analyses presented in Basu (2014). We study the distributional effects of alternative treatment modalities on seven-year health care expenditures among prostate cancer patients. Our data are derived from the 1995–2009 SEER-Medicare linked dataset. However, because of proprietary issues, we present an analysis on a simulated version of the original dataset, and this version is available with the `petiv` command. The key variables in our sample are categorized as the outcomes variable (Y), the treatment (D), the independent risk factors ($\mathbf{X}_O$), and the IV ($\mathbf{Z}$). These categories are common to any type of evaluation analysis.

(a) Outcomes variable (Y): Total undiscounted seven-year expenditures on health care expressed in 2009 dollars. Expenditures accumulate over all types of medical costs reimbursed by Medicare or a third-party payer and patients' out-of-pocket costs.

(b) Treatment (D): Comparison is made between the use of surgery in the first six months of diagnosis versus active surveillance that is defined as no use of surgery, hormone therapy, or radiation in the first six months of diagnosis along with at least two PSA tests within the first year of diagnosis. The treatment indicator takes a value of one for surgery.

(c) Independent risk factors ($\mathbf{X}_O$): These include clinical stage and grade of cancer for patients at diagnosis using standard definitions; demographics; an indicator for metropolitan area; Elixhauser comorbidity indices based on hospitalization in the year preceding diagnosis; year and state fixed effects; and zip-code level area characteristics on racial makeup, density, and education levels. We also adjust for hazard-rate ratios-level characteristics by using logged versions of population size, and per 100,000 patients' supply of hospital beds, physicians, specialists, and urologists.

(d) IV ($\mathbf{Z}$): hazard-rate ratios specific rates of active surveillance in prostate cancer patients in the year before the diagnosis of a patient.

Further details of this analysis can be found in Basu (2014). Variables were stored in the following macros:

```
. global xlist1 "age age2 T1 gradecat1 gradecat2 gradecat3 white black hispanic
> metro prehosp2 prehosp3 prehosp4"
. global xlist2 "chf valve  perivasc para neuro chrnlung dm dmcx hypothy obese
> lytes anemdef alcohol depress htn_c icd9_e_ch_mi icd9_e_ch_cevd icd9_e_ch_rd"
. global xlist3 "year2 year3 year4 year5 year6 year7 stateid2 stateid3 stateid4
> stateid5 stateid6 stateid7 stateid8 stateid9 stateid10 stateid11 stateid12"
. global xlist "$xlist1 $xlist2 $xlist3"
. global clist1 "xlist_zpnon00 xlist_zphso00 xlist_zpscl00 xlist_zpcol00
> xlist_zpblk00 xlist_zpwht00"
. global clist2 "ln_population ln_beds ln_physicians ln_specialists
> ln_urologists"
. global control "$clist1 $clist2"
. global iv "ivrate_activesurv"
. global y "payments_7years"
. global trt "surgery"
```

5.1 Applying the numerical algorithm to compute PeT effects

Our final analytic sample consisted of 13,495 patients, of whom 9,862 (73.3%) received surgery. We ran a logit model for the first stage where the indicator for surgery was regressed on $\mathbf{X}_O$ and $\mathbf{Z}$. The IV was found to be strongly predictive of surgery receipt conditional on other factors (F statistic: $10.9, p < 0.0001$). It was also found that the IV may be particularly suitable in reducing residual confounding in this application because it can reduce imbalance in observed factors considerably.

We then computed the predicted propensity score by using the standard `predict` command following the `logit` regression, as follows:

```
. use petanalysis_sjdata
. logit $trt $iv $xlist $control, robust
  (output omitted)
. predict ps, p
. matrix b = e(b)
. generate ps2= ps^2
. generate ps3=ps^3
. generate psr=round(ps,0.01)
. sort psr
```

```
. quietly forvalues ud=0(0.01)1 {
  2. local cnt1=0
  3. local cnt0=0
  4. tabulate $trt if round(psr, 0.01)==round(`ud´, 0.01) & $trt==1
  5. local cnt1=r(N)
  6. tabulate $trt if round(psr, 0.01)==round(`ud´, 0.01) & $trt==0
  7. local cnt0=r(N)
  8. noisily display "ud " `ud´ " cnt1 " `cnt1´ " cnt0 " `cnt0´
  9. if `cnt1´==0 |`cnt1´==. | `cnt0´==0 | `cnt0´==. {
 10.   noisily drop if round(psr, 0.01)==round(`ud´, 0.01)
 11.   noisily display " No coverage for P(Z) = " `ud´
 12.   }
 13. }
```

The identified support of the IV-based predicted propensity score, `ps`, which existed under both treatment arms, ranged from 0.07 to 0.998. Forty seven observations were lost because of lack of overlap. Thus our final analytic sample was 13,448.

We then determined appropriate specification for the LIV estimand. Because we are modeling health care expenditures, we started with a generalized linear model with log-link and gamma-variance specification. We included patient-specific covariates and their interactions with `ps`. We also included other supply-level variables as controls (but, to avoid overfitting, we did not include their interactions with `ps`). Finally, we varied the degree of polynomial for $\widehat{p}(\mathbf{x}, \mathbf{z})$ and tested alternatives using likelihood-ratio tests.

```
. global xlisti " "
. quietly foreach var of global xlist {
  2. capture drop p_`var´
  3. generate p_`var´ = ps*`var´
  4. global xlisti "$xlisti  p_`var´"
  5. }
. glm $y $xlisti $xlist $control ps, family(gamma) link(log) robust
  (output omitted)
. estimates store A
. glm $y $xlisti $xlist $control ps ps2, family(gamma) link(log) robust
  (output omitted)
. estimates store B
. lrtest A, force
Likelihood-ratio test                                 LR chi2(1)  =      0.14
(Assumption: A nested in B)                           Prob > chi2 =    0.7074
. glm $y $xlisti $xlist $control ps ps2 ps3, family(gamma) link(log) robust
  (output omitted)
. lrtest B, force
Likelihood-ratio test                                 LR chi2(1)  =      0.16
(Assumption: B nested in .)                           Prob > chi2 =    0.6904
```

Based on these results, the first-degree polynomial seemed most appropriate for these data. However, unlike linear models, a first-degree polynomial within a nonlinear model does not preclude the presence of essential heterogeneity in the additive scale (Heckman, Urzua, and Vytlacil 2006; Basu 2011).

Finally, we performed standard raw-scale residual-based goodness-of-fit tests with the generalized linear model with the chosen degree of polynomial for $\widehat{p}(\mathbf{x}, \mathbf{z})$ (Basu et al. 2007). No systematic biases were detected from residual-based goodness-of-fit analyses.

```
. glm $y $xlisti $xlist $control ps, family(gamma) link(log) robust
  (output omitted)
. predict xb, xb
. generate xb2=xb^2
. predict mu, mu
. generate res = $y - mu
. pwcorr res mu, sig
  (output omitted)
. glm $y xb xb2, family(gamma) link(log) robust
  (output omitted)
. xtile xbtile=xb, nq(20)
. tabulate xbtile, sum(res)
  (output omitted)
```

After completing the previous steps, we ran `petiv`.

```
. petiv $y $xlist, trt($trt) ps(ps) cmd(glm) controls($control) family(gamma)
> link(log) robust display
  (output omitted)
```

This created a variable called `pet_payments_7years` in the dataset. This variable may have missing values for a few observations where all MTEs in the range of the support of `ps` could not be calculated because of numerical overflow. In our case, `pet_payments_7years` was missing for 9 observations. Now mean treatment-effect parameters can be easily computed.

```
. // ATE
. summarize pet_payments_7years
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
pet_paymen~s |     13,441   -27427.69     43138.7  -550601.9   301051.5
. // TT
. summarize pet_payments_7years if $trt==1
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
pet_paymen~s |      9,856   -28674.51    46738.77  -550601.9   168978.2
. // TUT
. summarize pet_payments_7years if $trt==0
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
pet_paymen~s |      3,585    -23999.9     30914.8  -224579.6   301051.5
```

The ATE was estimated to be −$27,428. The effect on the treated and the effect on the untreated were −$28,675 and −$24,000, respectively. The distribution of individual treatment effects is illustrated using a histogram for the `pet_payments_7years` variable in figure 1. It is estimated that although the mean treatment-effect parameters are

negative, 22% of the patients are expected to incur positive expenditures from surgery as compared with that of watchful waiting over 7 years.

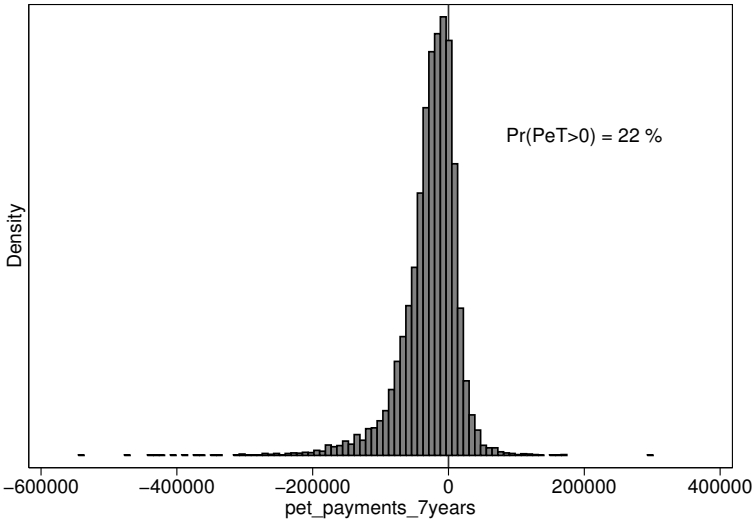

Figure 1. Distribution of `pet_payments_7years`

6 Conclusions

Here I have described a command, `petiv`, that uses an LIV approach to estimate PeT effects for a variety of specifications for the LIV estimand as outlined in Basu (2014). The `petiv` command creates a new variable in the dataset that contains the PeT effects for each individual in the dataset. However, the command takes the validity of the IVs and the specification of the LIV estimand as given. Appropriateness of these features of an LIV analysis should be determined before running the `petiv` command. The individual effects can be used to answer distributional questions and can also be easily aggregated to calculate mean treatment-effect parameters.

The estimator works best in analyses with larger sample sizes (say, over $N = 2000$). Such sample sizes are common in health economics and health policy applications. One also needs at least one continuous IV to estimate PeT effects. Finally, standard errors for PeT effects and the mean treatment effects can be obtained via bootstrap methods.

I hope that this methodology and the `petiv` command will be increasingly used in economics and other areas of research to help researchers understand distributional effects of interventions.

7 Acknowledgments

I would like to acknowledge support from the National Institute of Health Research Grants, RC4CA155809 and R01CA155329.

This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the author. I acknowledge the efforts of the Applied Research Program at the National Cancer Institute; the Centers for Medicare and Medicaid Services' Office of Research, Development, and Information; Information Management Services, Inc.; and the SEER Program tumor registries in the SEER-Medicare database.

The collection of the California cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract N01-PC-35136 awarded to the Northern California Cancer Center, contract N01-PC-35139 awarded to the University of Southern California, and contract N02-PC-15105 awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement #U55/CCR921930-02 awarded to the Public Health Institute. The ideas and opinions expressed herein are those of the author. Endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their contractors and subcontractors is not intended, nor should it be inferred.

8 References

Basu, A. 2011. Estimating decision-relevant comparative effects using instrumental variables. *Statistics in Biosciences* 3: 6–27.

———. 2014. Estimating person-centered treatment (PeT) effects using instrumental variables: An application to evaluating prostate cancer treatments. *Journal of Applied Econometrics* 29: 671–691.

Basu, A., J. J. Heckman, S. Navarro-Lozano, and S. Urzua. 2007. Use of instrumental variables in the presence of heterogeneity and self-selection: An application to treatments of breast cancer patients. *Health Economics* 16: 1133–1157.

Frandsen, B. R., M. Frölich, and B. Melly. 2012. Quantile treatment effects in the regression discontinuity design. *Journal of Econometrics* 168: 382–395.

Frölich, M., and B. Melly. 2013. Unconditional quantile treatment effects under endogeneity. *Journal of Business and Economic Statistics* 31: 346–357.

Heckman, J. J., S. Urzua, and E. J. Vytlacil. 2006. Understanding instrumental variables in models with essential heterogeneity. NBER Working Paper No. 12574, The National Bureau of Economic Research. http://www.nber.org/papers/w12574.

Heckman, J. J., and E. J. Vytlacil. 1999. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences* 96: 4730–4734.

———. 2001. Local instrumental variables. In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, ed. C. Hsiao, K. Morimune, and J. L. Powell, 1–46. New York: Cambridge University Press.

———. 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73: 669–738.

Imbens, G. W., and J. D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62: 467–475.

Newhouse, J. P., and M. McClellan. 1998. Econometrics in outcomes research: The use of instrumental variables. *Annual Review of Public Health* 19: 17–34.

Shaikh, A. M., and E. J. Vytlacil. 2011. Partial identification in triangular systems of equations with binary dependent variables. *Econometrica* 79: 949–955.

Vytlacil, E. 2002. Independence, monotonicity, and latent index models: An equivalence result. *Econometrica* 70: 331–341.

About the author

Anirban Basu is a professor in the Departments of Pharmacy, Health Services, and Economics at the University of Washington, where he directs the Program in Health Economics and Outcomes Methodology (PHEnOM). He is also a faculty research fellow at the National Bureau of Economic Research, Cambridge, MA.