Elasticities and the Inverse Hyperbolic Sine Transformation*

Marc F. Bellemare†
Casey J. Wichman‡

September 27, 2018

Abstract

Applied econometricians frequently apply the inverse hyperbolic sine (or arcsinh) transformation to a variable because it approximates the natural logarithm of that variable and allows retaining zero-valued observations. We provide derivations of elasticities in common applications of the inverse hyperbolic sine transformation and show empirically that the difference in elasticities driven by ad hoc transformations can be substantial. We conclude by offering practical guidance for applied researchers.

Keywords: Elasticities, Elasticity, Inverse Hyperbolic Sine, arcsinh
JEL Classification Codes: C1, C5

*We thank Jeff Bloem, Dave Giles, Seema Jayachandran, Jason Kerwin, David McKenzie, Steve Miller, and Joseph Ritter for useful comments and suggestions. All remaining errors are ours.

†Associate Professor, University of Minnesota, 1994 Buford Avenue, Saint Paul, MN 55108, mbellema@umn.edu

‡Fellow, Resources for the Future, 1616 P Street NW, Suite 600, Washington, DC 20036, wichman@rff.org.

1 Introduction

Applied econometricians often transform variables to make the interpretation of empirical results easier, to approximate a normal distribution, to reduce heteroskedasticity, or to reduce the effect of outliers. Taking the logarithm of a variable has long been a popular such transformation.

One problem with taking the logarithm of a variable is that it does not allow retaining zero-valued observations because $\ln(0)$ is undefined. But economic data often include meaningful zero-valued observations, and applied econometricians are typically loath to drop those observations for which the logarithm is undefined. Consequently, researchers have often resorted to ad hoc means of accounting for this when taking the natural logarithm of a variable, such as adding 1 to the variable prior to its transformation (MaCurdy and Pencavel, 1986).

In recent years, the inverse hyperbolic sine (or arcsinh) transformation has grown in popularity among applied econometricians because (i) it is similar to a logarithm, and (ii) it allows retaining zero-valued (and even negative-valued) observations (Burbidge et al., 1988; MacKinnon and Magee, 1990; Pence, 2006).[1]

For a random variable x, taking the inverse hyperbolic sine transformation

[1] As Ravallion (forthcoming) notes, in cases where a concave log-like transformation of negative data is necessary, the inverse hyperbolic sine transformation has undesirable properties, viz. it is convex over negative values. Given that, and because the inverse hyperbolic sine transformation is largely used with nonnegative variables (e.g., earnings), we focus in this paper only on cases where the variable to which the inverse hyperbolic sine transformation is applied is nonnegative.

yields a new variable $\tilde{x}$, such that

$$\tilde{x} = \text{arcsinh}(x) = \ln\left(x + \sqrt{x^2+1}\right). \tag{1}$$

The arcsinh transformation is easy to implement, yet there is no formal guidance on how to compute elasticities when using it, or on how it compares to alternative transformations.[2] In this article, we first fill this gap. We then provide an empirical illustration that relies on a widely known and publicly available data set. We conclude by offering guidelines for applied work.

2 Deriving Elasticities

In applied work, four common cases might present themselves to the applied econometrician, with each of the following pairs representing a dependent variable–explanatory variable combination: (1) linear–arcsinh, (2) arcsinh–linear, (3) arcsinh–linear with a dummy independent variable, and (4) arcsinh–arcsinh specifications. We derive elasticities analytically for each case. A useful preliminary result for what follows is that $\frac{\partial \tilde{x}}{\partial x} = \frac{1}{\sqrt{x^2+1}}$.

[2]For examples of the arcsinh transformation in practice, see Bahar and Rapoport (2018), Clemens and Tiongson (2017), Jayachandran et al. (2017), McKenzie (2017), or Muehlenbachs et al. (2017).

2.1 Linear–arcsinh Specification

Consider the following estimable equation

$$y = \alpha + \beta \tilde{x} + \epsilon. \tag{2}$$

In this case, computing $\widehat{\xi}_{yx} = \widehat{\frac{\partial y}{\partial x}} \cdot \frac{x}{y}$ after estimating the previous equation is relatively straightforward, since $\widehat{\frac{\partial y}{\partial x}} = \frac{\hat{\beta}}{\sqrt{x^2+1}}$, and so

$$\widehat{\xi}_{yx} = \frac{\hat{\beta}}{y} \frac{x}{\sqrt{x^2+1}}. \tag{3}$$

Because $\lim_{x \to \infty} \frac{x}{\sqrt{x^2+1}} = 1$, for large values of x, $\widehat{\xi}_{yx} \approx \frac{\hat{\beta}}{y}$.[3]

2.2 arcsinh–Linear Specification

Now, consider an estimable equation of the form

$$\tilde{y} = \alpha + \beta x + \epsilon \tag{4}$$

where x is a continuous variable. In this case, to recover y from the left-hand side of the equation of interest after estimating the previous equation, one has to apply the hyperbolic sine transformation—that is, the inverse of the

[3]We discuss what we mean by "large values of x" in sub-section 2.5.

inverse hyperbolic sine, or sinh—on both sides, so that

$$\hat{y} = \sinh(\hat{\alpha} + \hat{\beta}x + \hat{\epsilon}). \tag{5}$$

In this case, to recover $\widehat{\xi_{yx}} = \widehat{\frac{\partial y}{\partial x}} \cdot \frac{x}{y}$, we need to multiply $\frac{x}{y}$ by $\widehat{\frac{\partial y}{\partial x}} = \hat{\beta}\cosh(\hat{\alpha} + \hat{\beta}x + \hat{\epsilon})$ to get

$$\widehat{\xi_{yx}} = \hat{\beta}\cosh(\hat{\alpha} + \hat{\beta}x + \hat{\epsilon}) \cdot \frac{x}{y}, \tag{6}$$

which can be rewritten as

$$\widehat{\xi_{yx}} = \hat{\beta}\cosh(\text{arcsinh}(y)) \cdot \frac{x}{y} = \hat{\beta}x \cdot \frac{\sqrt{y^2+1}}{y}. \tag{7}$$

In this case, because $\lim_{y\to\infty} \frac{\sqrt{y^2+1}}{y} = 1$, for large values of y, $\widehat{\xi_{yx}} \approx \hat{\beta}x$.

2.3 arcsinh–Linear Specification with Dummy Independent Variables

Dummy variables pose a different set of problems for calculating elasticities, and this problem is not unique to arcsinh–linear econometric specifications. Standard log-linear specifications also suffer from this problem (Halvorsen and Palmquist, 1980; Kennedy, 1981). Consider the estimable equation

$$\tilde{y} = \alpha + \beta d + \epsilon \tag{8}$$

where d is a dichotomous variable taking on only values of zero or one. The quantity $\partial \tilde{y}/\partial d$ is undefined because dichotomous variables only change discretely. Following the arguments of Halvorsen and Palmquist (1980), Kennedy (1981), and Giles (1982) for semilogarithmic regression equations with dummy variables, we derive the analogous proportional effect for the arcsinh transformation, which we denote $\bar{P}$.

To derive $\bar{P}$, the percentage change in y associated with a change from $d = 0$ to $d = 1$, we first apply the sinh transformation to both sides, such that

$$\sinh(\tilde{y}) = y = \sinh(\alpha + \beta d + \epsilon). \tag{9}$$

After estimating Eq. 8, $\bar{P}$ is then defined as

$$\begin{aligned} \frac{\hat{\bar{P}}}{100} = \frac{\hat{y}(d=1) - \hat{y}(d=0)}{\hat{y}(d=0)} &= \frac{\sinh(\hat{\alpha} + \hat{\beta} + \hat{\epsilon}) - \sinh(\hat{\alpha} + \hat{\epsilon})}{\sinh(\hat{\alpha} + \hat{\epsilon})} \\ &= \frac{\sinh(\hat{\alpha} + \hat{\beta} + \hat{\epsilon})}{\sinh(\hat{\alpha} + \hat{\epsilon})} - 1 \end{aligned} \tag{10}$$

Researchers often adopt the arcsinh transformation to approximate a logarithm and interpret coefficients as they would for a logarithmic equation. In this case, one can approximate semi-elasticities as shown by Halvorsen and Palmquist (1980), Kennedy (1981), and Giles (1982) by exponentiating both sides of Eq. 8 and noting that

$$\exp(\tilde{y}) = y + \sqrt{y^2 + 1} \approx 2y \text{ for large } y.$$

The resulting approximation of a percentage change in y due to a discrete change in d is the standard Halvorsen and Palmquist (1980) result for logarithmic equations,[4]

$$\frac{\hat{\tilde{P}}}{100} \approx \exp(\hat{\beta}) - 1, \tag{11}$$

which, when using the small-sample bias correction suggested by Kennedy (1981), yields

$$\frac{\hat{\tilde{P}}}{100} \approx \exp(\hat{\beta} - 0.5\widehat{\mathrm{Var}}(\hat{\beta})) - 1, \tag{12}$$

where $\widehat{\mathrm{Var}}(\hat{\beta})$ is an estimate of the variance of $\widehat{\beta}$. For most applications, with large enough y, using $\hat{\tilde{P}}$ defined in Eq. 12 would suffice, rather than the exact interpretation, $\hat{\bar{P}}$. We note, however, that conditioning variables "drop out" of Eq. 12 due to the properties of natural logarithms, which is a special case. In general, it is important to include conditioning variables at their sample moments in elasticity formulations with non-logarithmic transformations, such as in Eq. 8.

2.4 arcsinh–arcsinh Specification

Now, consider a regression of the form,

$$\tilde{y} = \alpha + \beta\tilde{x} + \epsilon. \tag{13}$$

[4]See Appendix A for a derivation of this result.

In this last case, in order to recover y from the left-hand side of the equation of interest after estimating the previous equation, one has to apply the hyperbolic sine transformation on both sides, so that

$$y = \sinh(\hat{\alpha} + \hat{\beta}\tilde{x} + \hat{\epsilon}). \tag{14}$$

In this case, $\widehat{\frac{\partial y}{\partial x}} = \frac{\hat{\beta}\cosh(\hat{\alpha}+\hat{\beta}\text{arcsinh}(x)+\hat{\epsilon})}{\sqrt{x^2+1}}$, so that

$$\widehat{\xi_{yx}} = \frac{\hat{\beta}\cosh(\hat{\alpha} + \hat{\beta}\text{arcsinh}(x) + \hat{\epsilon})}{\sqrt{x^2+1}} \cdot \frac{x}{y}. \tag{15}$$

Because the interior of the cosh function is equal to $\tilde{y}$, we can rewrite this last equation as

$$\widehat{\xi_{yx}} = \frac{\hat{\beta}\cosh(\text{arcsinh}(y))}{\sqrt{x^2+1}} \cdot \frac{x}{y} = \hat{\beta} \cdot \frac{\sqrt{y^2+1}}{y} \cdot \frac{x}{\sqrt{x^2+1}}. \tag{16}$$

Because because $\lim_{x\to\infty} \frac{x}{\sqrt{x^2+1}} = 1$ and $\lim_{y\to\infty} \frac{\sqrt{y^2+1}}{y} = 1$, for large values of x and y, $\widehat{\xi_{yx}} \approx \hat{\beta}$.

2.5 Caveats

In the literature, the elasticities just derived are interpreted as the elasticities one would obtain from equivalent specifications with logarithmic transformations, but this only holds only for large enough average values of x, y, or both. This begs the question of how large is large enough. We suggest applied econometricians use approximate elasticities for values of x or y no less than 10. Although this value seems arbitrary, we suggest it because values greater than 10 minimize approximation error to less than half of a percent. Obviously, greater thresholds will yield greater accuracy.

That said, this is a rare situations where one can get something for (almost) nothing. Indeed, starting from the equation $y = \alpha + \beta x + \epsilon$, wherein $\xi_{yx} = \frac{\beta x}{y}$, multiplying y by $k > 0$ will return a coefficient on x equal to βk, and the elasticity ξ_{yx} will be such that $\frac{\partial y}{\partial x} \cdot \frac{x}{y} = \frac{\beta k x}{y k}$, so that $\xi_{yx} = \frac{\beta x}{y}$.[5] In this linear–linear case, rescaling y or x does not change the elasticity. But when an arcsinh transformation is involved, Monte Carlo simulations (not shown for brevity, but available upon request from the authors) show that ξ_{yx} only becomes more accurate as k increases, and that this holds for linear–arcsinh, arcsinh–linear, and arcsinh–arcsinh specifications. It is thus possible to obtain more accurate elasticities merely by multiplying the variable to which the arcsinh transformation is applied so as to deal with a variable whose mean is well above 10.

Lastly, recall that applied econometricians often adopt the arcsinh transformation to deal with variables with zero-valued observations, and the presence of such observations at low (i.e., zero) values of x or y may result in biased elasticity estimates. In these cases, researchers should calculate elasticities using their exact formulations, viz. those derived in equations 3, 7,

[5] Likewise, multiplying x by k will change the coefficient but leave the elasticity unaffected.

and 16. But because all three elasticities are undefined when $y = 0$, we suggest the following. Defining $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$, one workaround when $y = 0$ is to multiply by $\frac{\bar{x}}{\bar{y}}$ instead of by $\frac{x}{y}$, i.e., to report the elasticity at the means of x and y instead of one of the mean elasticities we derive above.

3 Empirical illustration

We illustrate common uses of the inverse-hyperbolic sine transformation using data from Dehejia and Wahba (1999), who evaluated LaLonde's (1986) canonical design-replication experiment for the National Supported Work (NSW) demonstration.

In that experiment, disadvantaged workers were randomly assigned either to the NSW program, wherein they were given counseling and work experience, or to the control group, wherein they received neither of those things. We use the Dehejia and Wahba (1999) subset of LaLonde's (1986) original data.

The outcome of interest is post-treatment annual earnings, Earnings_{78}. Because this application is for illustrative purposes, we focus on estimating the treatment effect of the NSW program controlling for pre-treatment earnings only. The inverse hyperbolic sine transformation of earnings is potentially attractive because earnings data tends to be right-skewed and there is a nontrivial number of true zeros in the data. As shown in Table 1, Earnings_{78} has 45 and 92 zero-valued observations respectively for the treatment and control groups, with 111 and 178 zero-valued pre-treatment earnings observations respectively for the treatment and control groups.

We estimate four simple specifications that capture each of the cases described above, as well as two models with the oft-used $\ln(x+1)$ transformation for comparison. Results are presented in Table 2.

In column (1), we see that the NSW job-training program increased post-treatment earnings of the treated group by $1,750 per year relative to the control group. This estimate is illustrative of the estimate in Dehejia and Wahba (1999) Table 3, panel B, row 1. We calculate a semi-elasticity by dividing the treatment coefficient by the sample mean of the dependent variable to show that NSW increased earnings of treated individuals by approximately 33%. We also calculate the earnings elasticity between pre- and post-treatment earnings, which shows that as pre-treatment earnings increase by 10%, we observe a 4.3% increase in post-treatment earnings. This elasticity does not have a meaningful economic interpretation, although it is useful for comparing elasticity calculations.

In column (2), we transform pre-treatment earnings using the arcsinh transformation. This adjustment reduces the earnings elasticity from 0.043 to 0.026, a nearly 40 percent reduction. This is a drastic change for a seemingly innocuous transformation of an independent variable, likely due to having reduced the influence of outliers (e.g., very high earners on average) as a result of the inverse hyperbolic sine transformation. The coefficient on NSW, however, remains stable, suggesting that the arcsinh transformation of a control variable does not affect causal interpretation of treatment variables and can improve the fit of the model.

In columns (3) and (4), we repeat this analysis for arcsinh transformation of an independent variable: post-treatment earnings. Here, we see notable deviations from the linear–arcsinh case. Focusing on the earnings elasticity, we find a similar pattern as in columns (1) and (2): applying the arcsinh transformation to post-treatment earnings reduces the elasticity estimate by a factor of two. These elasticities, overall, are substantially larger than that of the linear case. This illustration shows that a transformation of the dependent variable can have a drastic effect on elasticity estimates.

We also present in columns (3) and (4) semi-elasticity estimates of NSW on earnings calculated both by exact methods ($\bar{P}$) and by the logarithmic approximation ($\tilde{P}$). Results suggest that NSW increased earnings by 158–188 percent. These results highlight the danger of interpreting coefficient estimates directly as semi-elasticities for large coefficients on dummy variables in arcsinh–linear regression equations. In both columns, $\bar{P}$ is approximately 25 percentage points larger than $\tilde{P}$. This difference stems almost entirely from the small-sample adjustment in $\tilde{P}$. Because earnings are large in our application, the difference in $\bar{P}$ and $\tilde{P}$ before the small-sample adjustment is trivial, suggesting the simpler logarithmic approximation performs well.

Finally, in columns (5) and (6), we repeat the previous analysis but substitute the $\ln(x+1)$ transformation for the arcsinh transformation. Our results for $\xi(\text{Earnings}_{78}, \text{Earnings}_{75})$ are largely analogous to the arcsinh case. This similarity suggests that common transformations of dependent and independent variables with a large proportion of zeros and outliers can have substantial implications for elasticity estimates. For semi-elasticities, we find that transforming dependent variables by logarithms produces percentage changes about 15–20 percentage points smaller than arcsinh transformations (shown in columns (3) and (4)) in our application.

We do not take a stand on whether the effect of the NSW program presented in this sequence of results is the true effect of job-training programs on earnings; that is not the point of this paper. Rather, we show that under seemingly innocuous, oft-used transformations of dependent and independent variables, we can produce wildly different elasticity estimates even within a randomized design. We view this simple analysis as a cautionary warning to applied econometricians about using ad hoc transformations to deal with zero-valued observations and to facilitate easy interpretation of empirical results.

4 Summary and Concluding Remarks

We have first derived exact elasticities in cases where an applied econometrician applies the popular inverse hyperbolic sine (or arcsinh) transformation to a variable, characterizing elasticities for the cases where the arcsinh transformation is applied to (i) an explanatory variable of interest, (ii) the dependent variable with a continuous explanatory variable of interest, (iii) the dependent variable with a dichotomous explanatory variable of interest, and (iv) both the dependent variable and the explanatory variable of interest.

After discussing some of the caveats of our approach, we have derived those elasticities for a well-known application, which has provided a cautionary tale regarding the use of ad hoc transformations of variables when dealing with zero-valued observations.

We conclude with the following guidelines for applied researchers who wish to use the inverse hyperbolic sine transformation in their own work to obtain elasticities:

1. One needs to be careful when interpreting inverse hyperbolic sine coefficients as semi-elasticities. In most cases, the researcher will need to transform those coefficients in the manner derived above prior to interpreting them as percentage changes. In our empirical example, interpreting β on NSW as a semi-elasticity in an arcsinh–linear equation understates the correct percentage effect by 40 percent.

2. Standard logarithmic adjustments for semi-elasticities in arcsinh–linear equations with dummy variables can be used with little error for dependent variables with untransformed means roughly greater than 10.

3. Applying the inverse hyperbolic sine transformation to an explanatory variable of interest appears somewhat harmless, but it can change elasticities relative to the linear-linear model substantially.

4. For cases where the applied researcher is dealing with data that contain many zero-valued observations, it is probably better to model the data-generating process explicitly, e.g., using a Tobit, negative binomial, or zero-inflated regression model.

5. $\text{arcsinh}(x) \neq \ln(x)$ for small values of x, so in cases where the applied econometrician is interested in using the inverse hyperbolic sine transformation to study a change from $x = 0$ to a positive but small value (e.g., $x < 10$), the inverse hyperbolic sine may not be the right transformation.

6. It is easy to obtain a more accurate elasticity estimate by multiplying the variable to which the arcsinh transformation is applied by a constant $k > 1$. As k increases, the elasticity estimate becomes more accurate. So for example, one can get more accurate estimates by taking the arcsinh of a price in dollars rather than of the same price in thousands of dollars.

Overall, the inverse hyperbolic sine function can be a useful tool for econometricians using variables with extreme values and true zeros. For large positive values, arcsinh can generally be treated like a natural logarithm. Future research might consider small-sample adjustments in the elasticity formulations we have presented here.

Table 1: Summary statistics of Dehejia and Wahba (1999) NSW data

	Treatment Group		Control Group	
	Mean	SD	Mean	SD
Earnings_{78}	6349.1	7867.4	4554.8	5483.8
Earnings_{75}	1532.1	3219.3	1266.9	3103.0
Observations	185		260	
$\text{Earnings}_{78} = 0$	45		92	
$\text{Earnings}_{75} = 0$	111		178	

Notes: This table presents summary statistics for the Dehejia and Wahba (1999) sample of LaLonde's (1986) original data. Only earnings data and treatment indicators are used in this analysis.

Table 2: Illustration of common inverse-hyperbolic sine transformations and elasticities using data from Dehejia and Wahba (1999)

	Earnings_{78}		$\text{arcsinh}(\text{Earnings}_{78})$		$\ln(\text{Earnings}_{78}+1)$	
	(1)	(2)	(3)	(4)	(5)	(6)
NSW	1750.2***	1694.3***	1.056**	1.034**	0.982**	0.960**
	(632.1)	(633.9)	(0.416)	(0.417)	(0.386)	(0.387)
Earnings_{75}	0.167*		0.000112*		0.000105*	
	(0.0990)		(0.0000651)		(0.0000604)	
$\text{arcsinh}(\text{Earnings}_{75})$		135.5*		0.0705		
		(76.86)		(0.0506)		
$\ln(\text{Earnings}_{75}+1)$						0.0722
						(0.0510)
Constant	4343.6***	4197.8***	5.803***	5.759***	5.365***	5.323***
	(426.1)	(454.7)	(0.280)	(0.299)	(0.260)	(0.278)
Adjusted R^2	0.0197	0.0203	0.0173	0.0150	0.0174	0.0152
Observations	445	445	445	445	445	445
Earnings_{78} (treatment mean)	6349.1	6349.1	6349.1	6349.1	6349.1	6349.1
Earnings_{78} (control mean)	4554.8	4554.8	4554.8	4554.8	4554.8	4554.8
Earnings_{75} (treatment mean)	1532.1	1532.1	1532.1	1532.1	1532.1	1532.1
Earnings_{75} (control mean)	1266.9	1266.9	1266.9	1266.9	1266.9	1266.9
Calculated (semi-)elasticities:						
$\xi(\text{Earnings}_{78},\ \text{NSW})$	0.330	0.320				
$\xi(\text{Earnings}_{78},\ \text{Earnings}_{75})$	0.0433	0.0256	0.155	0.0705	0.144	0.0722
$\tilde{P}(\text{Earnings}_{78},\ \text{NSW})/100$			1.638	1.578	1.477	1.424
$\bar{P}(\text{Earnings}_{78},\ \text{NSW})/100$			1.876	1.813		

Notes: This table presents illustrative regression results using data from the evaluation of the National Supported Work experiment by Dehejia and Wahba (1999) and reproduced in Table 3.3.3 of Angrist and Pischke (2008). Data are obtained from `https://economics.mit.edu/faculty/angrist/data1/mhe/dehejia` (Last Accessed: June 1, 2018). NSW is a dummy variable signifying assignment to the randomized treatment group. Earnings_{78} and Earnings_{75} are annual earnings (in dollars) for 1978 and 1975, respectively. All (semi-)elasticities are calculated at the sample mean.

References

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton university press, 2008.

Bahar, Dany and Hillel Rapoport, "Migration, Knowledge Diffusion, and the Comparative Advantage of Nations," *Economic Journal*, 2018, *128* (July), F273–F305.

Burbidge, John B., Lonnie Magee, and A. Leslie Robb, "Alternative transformations to handle extreme values of the dependent variable," *Journal of the American Statistical Association*, 1988, *83* (401), 123–127.

Clemens, Michael A. and Erwin R. Tiongson, "Split Decisions: Household Finance When a Policy Discontinuity Allocates Overseas Work," *Review of Economics and Statistics*, 2017, *99* (3), 531–543.

Dehejia, Rajeev H. and Sadek Wahba, "Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs," *Journal of the American Statistical Association*, 1999, *94* (448), 1053–1062.

Giles, David E.A., "The interpretation of dummy variables in semilogarithmic equations: Unbiased estimation," *Economics Letters*, 1982, *1-2* (10), 77–79.

Halvorsen, Robert and Raymond Palmquist, "The interpretation of dummy variables in semilogarithmic equations," *American Economic Review*, 1980, *70* (3), 474–75.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas, "Cash for Carbon: A Randomized Trial of Payments for Ecosystem Services to Reduce Deforestation," *Science*, 2017, *21* (6348), 267–273.

Kennedy, Peter E., "Estimation with correctly interpreted dummy variables in semilogarithmic equations," *American Economic Review*, 1981, *71* (4).

LaLonde, Robert J., "Evaluating the econometric evaluations of training programs with experimental data," *American Economic Review*, 1986, pp. 604–620.

MacKinnon, James G. and Lonnie Magee, "Transforming the dependent variable in regression models," *International Economic Review*, 1990, pp. 315–339.

MaCurdy, Thomas E. and John H. Pencavel, "Testing between competing models of wage and employment determination in unionized markets," *Journal of Political Economy*, 1986, *94* (3, Part 2), S3–S39.

McKenzie, David, "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition," *American Economic Review*, 2017, *107* (8), 2278–2307.

Muehlenbachs, Lucija, Stefan Staubli, and Ziyan Chu, "The Accident Externality from Trucking," Technical Report, National Bureau of Economic Research 2017.

Pence, Karen M., "The role of wealth transformations: An application to estimating the effect of tax incentives on saving," *The B.E. Journal of Economic Analysis & Policy*, 2006, *5* (1).

Ravallion, Martin, "A Concave Log-Like Transformation Allowing Non-Positive Values," *Economics Letters*, forthcoming.

Appendix A

Consider the model

$$\text{arcsinh}(y) = \alpha + \beta d + \epsilon \tag{A.1}$$

To approximate $\bar{P}$, the percentage change in $\hat{y}$ due to a change in the dummy variable d, we exponentiate both sides of Eq. A.1 after estimation,

$$\exp(\text{arcsinh}(\hat{y})) = \hat{y} + \sqrt{(}\hat{y}^2 + 1) \approx 2\hat{y} = \exp(\hat{\alpha} + \hat{\beta}x + \hat{\epsilon}) \tag{A.2}$$

where the approximation holds for large $\hat{y}$.

We then define the percentage change in $\hat{y}$ due to a discrete change in d,

$$\begin{aligned} \frac{\bar{P}}{100} =& \frac{\hat{y}(d=1) - \hat{y}(d=0)}{\hat{y}(d=0)} \\ =& \frac{\frac{1}{2}\exp(\hat{\alpha} + \hat{\beta}x + \hat{\epsilon}) - \frac{1}{2}\exp(\hat{\alpha} + \hat{\epsilon})}{\frac{1}{2}\exp(\hat{\alpha} + \hat{\epsilon})} = \exp(\hat{\beta}) - 1. \end{aligned} \tag{A.3}$$