



University of Victoria

Department of Economics

Econometrics Working Paper EWP1106

ISSN 1485-6441

## On the Inconsistency of Instrumental Variables Estimators for the Coefficients of Certain Dummy Variables

David E. Giles

*Department of Economics, University of Victoria  
Victoria, B.C., Canada V8W 2Y2*

May, 2011

### Abstract

In this paper we consider the asymptotic properties of the Instrumental Variables (IV) estimator of the parameters in a linear regression model with some random regressors, and other regressors that are dummy variables. The latter have the special property that the number of non-zero values is fixed, and does not increase with the sample size. We prove that the IV estimator of the coefficient vector for the dummy variables is *inconsistent*, while that for the other regressors is weakly consistent under standard assumptions. However, the usual estimator for the asymptotic covariance matrix of the I.V. estimator for *all* of the coefficients retains its usual consistency. The t-test statistics for the dummy variable coefficients are still asymptotically standard normal, despite the inconsistency of the associated IV coefficient estimator. These results extend the earlier results of Hendry and Santos (2005), which relate to a fixed-regressor model, in which the dummy variables are non-zero for just a single observation, and OLS estimation is used.

**Keywords**                      Dummy variables; indicator variables; instrumental variables; inconsistent estimator

**JEL Classifications**        C20; C26; C49

**Word Count**                    2,948

---

### Author Contact:

David E. Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: dgiles@uvic.ca; Phone: (250) 721-8540; FAX: (250) 721-6214

## 1. Introduction

Dummy (“indicator”) variables are routinely used as covariates in regression analysis. These binary variables are assigned the value zero if the corresponding sample observation fails to exhibit a particular qualitative attribute, and a constant non-zero value if that attribute is achieved. Without loss of generality, this single value is assumed to be unity. For any other value, the corresponding regression coefficients simply scale to compensate. It is well known (Salkever, 1976) that if a dummy variable regressor takes its non-zero value for only a single sample value, then the OLS estimates of the regression coefficients are identical to those that would be obtained if the dummy variable were omitted from the model, and the observation in question were deleted from the sample for all of the model’s variables. This result should be kept in mind if such dummy variables are used to deal with “outliers” in the data.

Recently, Hendry and Santos (2005) have discussed another intriguing property of OLS estimation in the case of such special dummy variables. Namely, even under standard assumptions, the OLS estimator of the coefficient of such a dummy variable is inconsistent, even though OLS is still best linear unbiased. In contrast, the OLS estimator for the coefficient vector associated with the remaining “regular” regressors retains its usual weak consistency property. Hendry and Santos provide other related results that are of considerable practical importance.

In this paper we extend the principal results of Hendry and Santos in two directions. First, we show that corresponding results hold when the non-dummy regressors in the model are stochastic, and correlated with the error term, and OLS is replaced by Instrumental Variables (IV) estimation. Second, we observe that the basic results, and their proofs, are essentially unchanged if the problematic dummy variables have a non-zero constant value for a fixed number of observations, rather for a single observation, provided that this number cannot increase as the sample size increases.

Models with these characteristics are commonly encountered in applications involving both time-series and cross-section data. In the former case, one obvious example would be where the sample spans the years of the Second World War, and a dummy variable is assigned to be non-zero only for the seven years of conflict. A second time-series example would be where the sample spans the “9/11” terrorist attacks in the U.S.. In this case, it is assumed that the earliest recorded data are included in the sample already, and that any future such catastrophes would be

of a sufficiently different nature that separate dummy variables would be required if the sample were extended into the future. In the case of cross-section data, the situation under consideration here can arise in the analysis of treatment effects using a “differences-in-differences” methodology. Here, dummy variables with a fixed number of non-zero observations will arise, for example with respect to the pre- and post-treatment periods. Whatever the type of data being used, dummy variables with only a single non-zero value are often used to deal with outlier values in the sample, as noted already.

The widespread use of IV estimators, in both time-series and cross-sectional data analyses, means that it is important to understand their properties when they are applied in the context of these “special” dummy variable regressors. Moreover, in many applications, the inferences that are of primary interest centre on the coefficients of the dummy variables themselves. For example, one might be concerned with the impact of a structural break; and in the case of differences-in-differences estimation, attention is focused on the coefficient of the product of two dummy variables, each of which could be of the form under consideration here. If the estimates of these coefficients are inconsistent, this severely weakens any conclusions that may be drawn from the study.

The rest of the paper is structured as follows. In the next section we formalize the model and the underlying assumptions that we will use. Apart from the characteristics of the dummy variables, everything is quite standard. In sections 3 and 4 we establish our main results. These include the consistency of the IV estimator for the coefficients of the non-dummy regressors; the *inconsistency* of this estimator in the case of the coefficients of the dummy variables themselves; and the consistency of the usual estimator for the asymptotic covariance matrix of the IV estimator of *all* of the coefficients in the model. In Section 5 we demonstrate that although the IV estimator for the coefficients of the special dummy variables is inconsistent, the associated t-statistics are still asymptotically standard normally distributed, as usual. So, asymptotically valid inferences may still be drawn, although the t-test is an inconsistent test in this context and so its power is limited, even asymptotically. Section 6 provides our conclusions.

## 2. Model and assumptions

Consider the following linear regression model:

$$y = X\beta + D\gamma + v \quad ; \quad v \sim N[0, \sigma^2 I] \quad (1)$$

where  $X$  is  $(n \times k_1)$ ,  $D$  is  $(n \times k_2)$  and  $k = k_1 + k_2$  is the total number of regressors, including the intercept. At least some of the columns of  $X$  are random and correlated even asymptotically with the error term,  $v$ . That is  $plim(n^{-1}X'v) \neq 0$ . The columns of  $D$  are zero-one indicator variables, each taking the value unity only for one (different) observation. As we will see in the next section, subject to  $X$  having full rank, we can generalize this set-up to allow each dummy variable to take the value unity only for a fixed number of observations that does not increase with the sample size. All of our results hold under this more general framework. In addition our results are unaffected if the number of zero values for the dummy variable(s) is fixed, but the number of unit values can increase. If  $\mathbf{1}$  is a unit-valued vector, replacing a column,  $d$ , of  $D$  with  $(\mathbf{1} - d)$  simply reverses the sign of the corresponding element of  $\gamma$ .

Given the presence of stochastic regressors, instrumental variables (IV) estimation would be a natural choice in anticipation of obtaining weakly consistent estimates of the parameters. However, as we will show, this hope is thwarted as far as the estimation of the elements of  $\gamma$  are concerned. Without loss of generality, we will consider the just-identified case, with equal numbers of regressors and instruments. Also without loss of generality we will include the intercept and all of the columns of  $D$  in the set of instruments, unless otherwise noted. Let the columns of  $Z$  ( $n \times k_1$ ) be the remaining  $k_1$  instruments, satisfying the conditions  $plim(n^{-1}Z'v) = 0$  and  $plim(n^{-1}Z'X) = Q_{ZX}$ , where  $Q_{ZX}$  is finite and non-singular.

For expository purposes, we let  $k_2 = 1$ , so that  $D$  is a single column vector,  $d$ , with one non-zero element at observation,  $i_b$ , say. Note that  $d'd = 1$ ,  $X'd = x_{i_b}$  and  $Z'd = z_{i_b}$  where  $x_{i_b}$  and  $z_{i_b}$  are  $(k_1 \times 1)$  vectors with elements comprising the values of the regressors in  $X$  and the instruments in  $Z$ , respectively, at observation  $i_b$ . So,  $plim(n^{-1}d'X) = 0'$ . Also, define  $M_1 = I - X(Z'X)^{-1}Z'$  and  $M_2 = I - d(d'd)^{-1}d' = I - dd'$ , so that  $M_1X = 0$  and  $M_2d = 0$ . In addition, note that

$$plim(d' M_1 d) = 1 - plim[(n^{-1} d' X)(n^{-1} Z' X)^{-1} Z' d] = [1 - 0' Q_{ZX}^{-1} z_{i_b}] = 1. \quad (2)$$

We can now proceed to our main results.

### 3. (In)Consistency of the IV estimators

First, note the following obvious extension of the Frisch-Waugh-Lovell theorem (Frisch and Waugh, 1933; Lovell 1963, 2008) to our situation involving IV estimation. In the notation of our model, the IV estimators of  $\beta$  and  $\gamma$  are  $\tilde{\beta} = (Z' M_2 X)^{-1} Z' M_2 y$  and  $\tilde{\gamma} = (d' M_1 d)^{-1} d' M_1 y$ . This is easily established by solving the two “normal equations”:

$$\begin{aligned} Z' X \beta + Z' d \gamma &= Z' y \\ d' X \beta + d' d \gamma &= d' y \end{aligned} \quad (3)$$

for  $\beta$  and  $\gamma$ . See Giles (1984) for a related application of this result.

We now show that the IV estimator of  $\gamma$  is inconsistent, but that of  $\beta$  is weakly consistent under standard assumptions.

#### **Theorem 1**

For model (1) and the assumptions in section 2,

$$plim(\tilde{\gamma} - \gamma) = v_{i_b} \neq 0.$$

#### **Proof**

$$\begin{aligned} \tilde{\gamma} &= (d' M_1 d)^{-1} d' M_1 y \\ &= (d' M_1 d)^{-1} d' M_1 (X \beta + d \gamma + v) \end{aligned}$$

Using the result,  $M_1 X = 0$ , we have:

$$\begin{aligned} \tilde{\gamma} - \gamma &= (d' M_1 d)^{-1} d' M_1 v \\ &= (d' M_1 d)^{-1} d' v - (d' M_1 d)^{-1} d' X (Z' X)^{-1} Z' v \end{aligned} \quad (4)$$

Note, from (2), and Slutsky's Theorem, that

$$plim[(d' M_1 d)^{-1} (d' X)(n^{-1} Z' X)^{-1} (n^{-1} Z' v)] = (x_{i_b} Q_{ZX}^{-1} 0) = 0. \quad (5)$$

Also, recalling that  $d' M_1 d = 1$ ,

$$plim[(d' M_1 d)^{-1} (d' v)] = plim(v_{i_b}) = v_{i_b} ,$$

where  $v_{i_b}$  is a single element of  $v$ .

So,

$$plim(\tilde{\gamma} - \gamma) = v_{i_b} \neq 0 .$$

■

Note that if the dummy variable takes the value unity for a *fixed* number of observations (the first  $m$ , say), and this number does not increase with  $n$ , then  $\tilde{\gamma}$  is still inconsistent. Specifically, following the above proof, in this case  $plim(d' M_1 d) = m$ ;  $plim[(d' M_1 d)^{-1} d' X (Z' X)^{-1} Z' v] = 0$ ; and  $plim[(d' M_1 d)^{-1} (d' v)] = m^{-1} plim(\sum_{j=1}^m v_j) = (\sum_{j=1}^m v_j) / m = \bar{v}_{(m)}$ , say. It then follows that  $plim(\tilde{\gamma} - \gamma) = \bar{v}_{(m)} \neq 0$ .

Also, note that the result in Theorem 1 does not require that the dummy variable(s) be included in the set of instruments. Using similar arguments it is easily seen that  $\tilde{\gamma}$  is still inconsistent for  $\gamma$  if the matrix of instruments is  $[Z, w]$ , where the random vector  $w$  satisfies  $plim(n^{-1} w' v) = 0$ . Parenthetically, one might think of “adjusting” the IV estimator by subtracting the  $i_b^{th}$  IV residual from  $\tilde{\gamma}$ , in the case of a dummy with a single non-zero value, in anticipation of obtaining a consistent estimator of  $\gamma$ . This is futile, as the residual in question has a zero value. The same applies when there are  $m$  non-zero observation. In that case the sum (and sample average) of the corresponding IV residuals is also zero.

Now consider the IV estimator of  $\beta$ , the coefficient vector for the random regressor matrix,  $X$ .

### **Theorem 2**

For model (1) and the assumptions in section 2,  $\tilde{\beta}$  is weakly consistent for  $\beta$ .

### **Proof**

$$\begin{aligned}\tilde{\beta} &= (Z' M_2 X)^{-1} Z' M_2 y \\ &= (Z' M_2 X)^{-1} Z' M_2 (X\beta + d\gamma + v) .\end{aligned}$$

Using the results,  $M_2 d = 0$  and  $d' d = 1$ , we have:

$$\begin{aligned}\tilde{\beta} - \beta &= (Z' M_2 X)^{-1} Z' M_2 v \\ &= (Z' M_2 X)^{-1} Z' v - (Z' M_2 X)^{-1} Z' d d' v .\end{aligned}\tag{6}$$

Applying Slutsky's Theorem repeatedly,

$$\begin{aligned}plim(n^{-1} Z' M_2 X) &= plim(n^{-1} Z' X) - plim(n^{-1} z_{i_b} x_{i_b}') \\ &= Q_{ZX} .\end{aligned}\tag{7}$$

So,

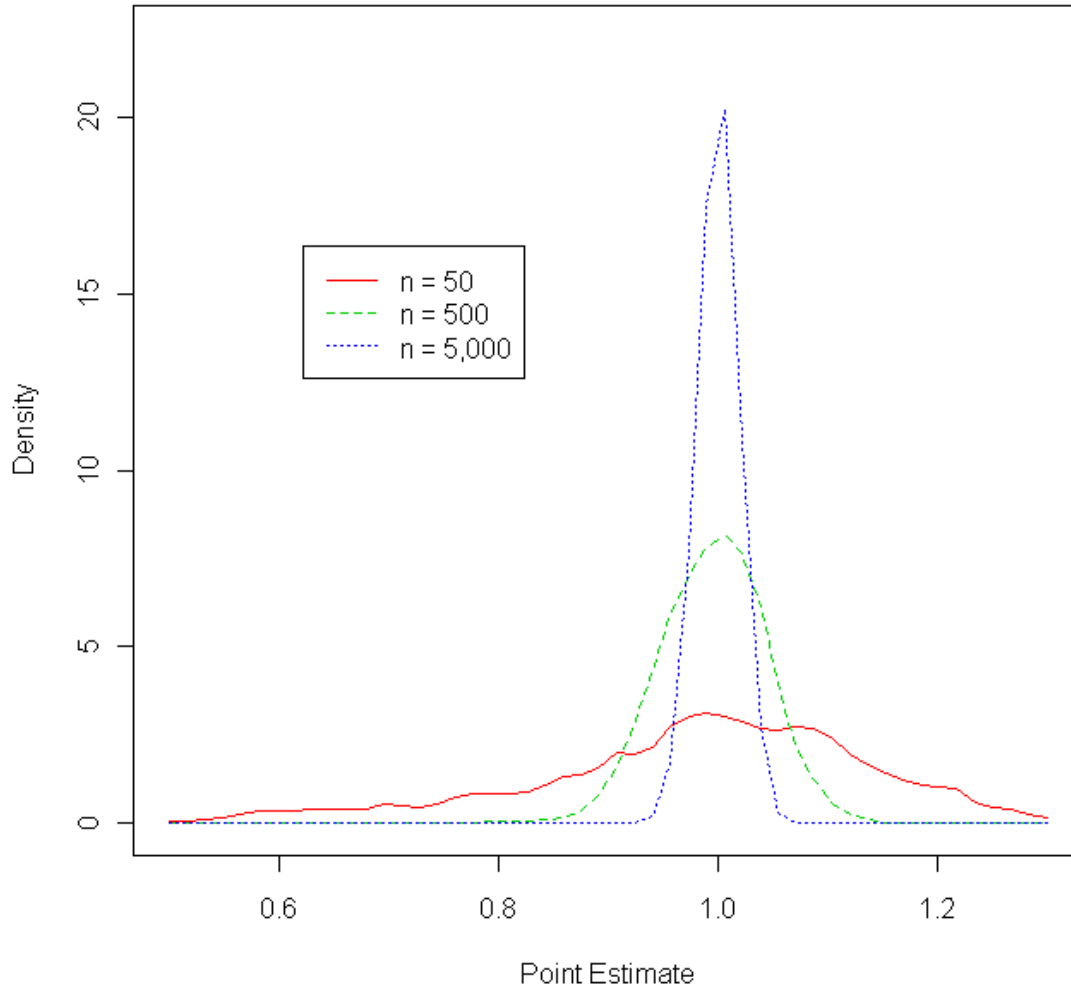
$$\begin{aligned}plim[(Z' M_2 X)^{-1} (Z' d)(d' v)] &= Q_{ZX}^{-1} plim[(n^{-1} Z' d)(d' v)] \\ &= (Q_{ZX}^{-1} 0 v_{i_b}) \\ &= 0 , \\ plim[(Z' M_2 X)^{-1} (Z' v)] &= Q_{ZX}^{-1} plim(n^{-1} Z' v) \\ &= (Q_{ZX}^{-1} 0) \\ &= 0 ,\end{aligned}$$

and therefore  $plim(\tilde{\beta} - \beta) = 0$  .

■

So, the IV estimator for the coefficient of the non-dummy variable is still weakly consistent, despite the degenerate nature of the dummy variable itself. Again, this result also holds if the dummy variable takes the value unity for a *fixed* number of observations, and this number does not increase with  $n$ . It also holds in the case of several such dummy variables. This is easily established by making obvious amendments to the proof of Theorem 2, and is left as an exercise for the interested reader.

**Figure 1: Sampling Distributions - X Coefficient Estimator**



To illustrate some of the above results the sampling distributions of  $\tilde{\beta}$  and  $\tilde{\gamma}$  have been simulated, using program code written for the EViews 7.1 package (Quantitative Micro Software, 2010). The experiment is based on the following data-generating process:

$$x_i = a_0 + a_1 d_i + a_2 z_i + \varepsilon_i \quad (8)$$

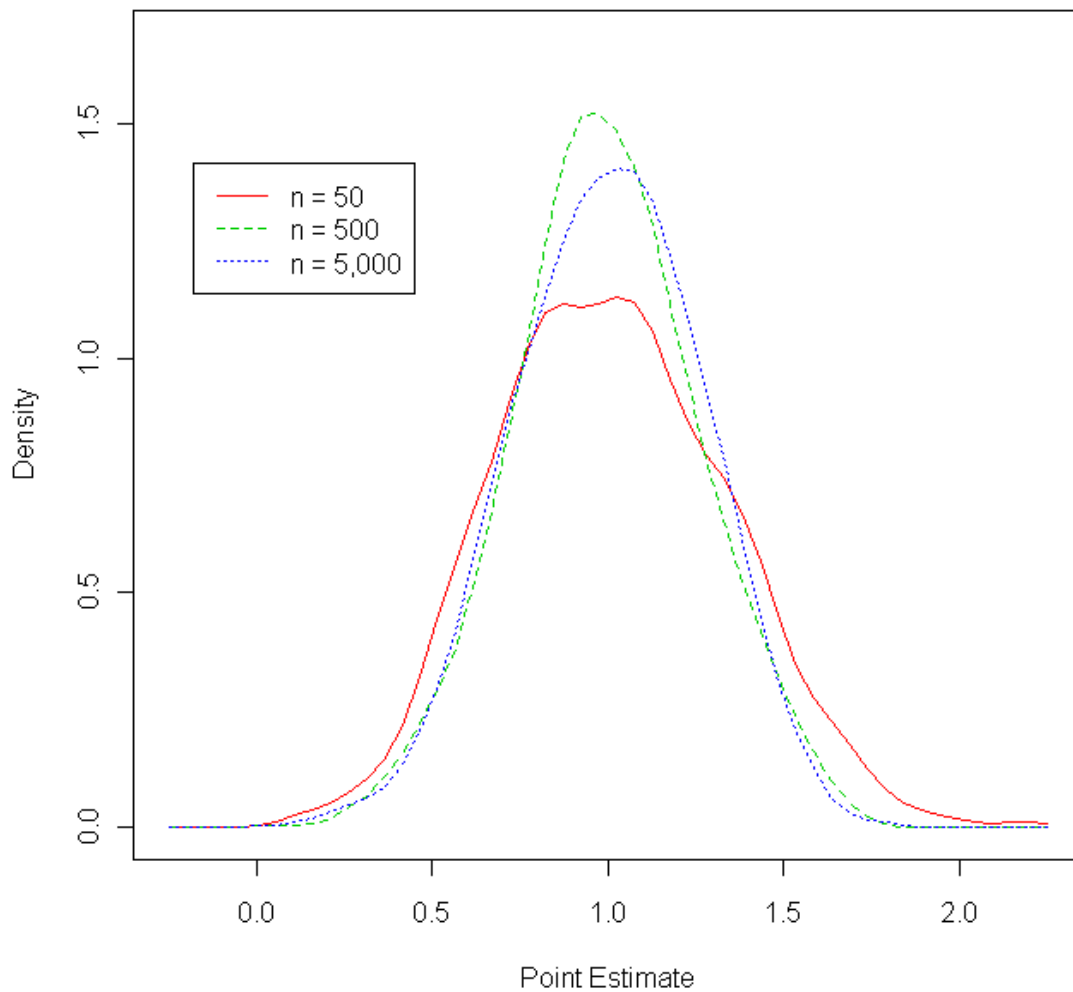
$$y_i = \alpha + \beta x_i + \gamma d_i + v_i \quad (9)$$

$$\begin{pmatrix} \varepsilon_i \\ v_i \end{pmatrix} \sim N \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} ; \quad i = 1, 2, \dots, n.$$



Sample sizes up to  $n = 200,000$  were considered. The exogenous variable,  $z$ , was generated as standard normal for the largest sample size, and held fixed for the 1,000 Monte Carlo repetitions. The dummy variable,  $d$ , was constructed as  $d_i = 1, i = 1, 2, \dots, 15; d_i = 0, i = 16, 17, \dots, n$ . The following parameter values were assigned:  $a_0 = a_1 = a_2 = \alpha = \beta = \gamma = \sigma_1^2 = \sigma_2^2 = 1; \sigma_{12} = 0.95$ . Equation (9) was estimated by IV with  $d, z$  and the intercept as instruments. The high correlation between the regressor,  $x$ , and the instrument,  $z$ , ensures that we avoid any “weak instruments” problem. The effective signal-to-noise ratios implies that the median  $R^2$  values for the 1,000 IV regressions were 0.818, 0.795, and 0.796 for  $n = 50, 500$  and 5,000 respectively. Figures 1 and 2 illustrate the sampling distributions of  $\tilde{\beta}$  and  $\tilde{\gamma}$  for a selection of sample sizes.

**Figure 2: Sampling Distributions - Dummy Coefficient Estimator**



The weak consistency of  $\tilde{\beta}$  is apparent in Figure 1. In contrast, Figure 2 demonstrates the inconsistency of  $\tilde{\gamma}$ , and the fact that this arises because the variance of that estimator does not vanish as  $n \rightarrow \infty$ .

#### 4. Asymptotic variances of the IV estimators

Hendry and Santos (2005, pp. 573-574) show that although the OLS estimator of  $\gamma$  is inconsistent in their fixed-regressor, single-valued dummy variable analogue to our problem, the usual estimator of the variance of that OLS estimator is both unbiased and consistent. In our own case, we will now show that the usual estimator of the asymptotic variance for  $\tilde{\gamma}$  is weakly consistent, notwithstanding the inconsistency of  $\tilde{\gamma}$  itself.

##### **Theorem 3**

Consider model (1) and the assumptions in section 2. In addition, assume that  $plim(n^{-1}X'v)$  is a finite vector; and  $plim(n^{-1}X'X) = Q_{XX}$  and  $plim(n^{-1}Z'Z) = Q_{ZZ}$  are finite matrices. Then the asymptotic variance of  $\sqrt{n}(\tilde{\gamma} - \gamma - v_{i_b})$  is  $(\sigma^2 x_{i_b}' Q_{ZX}^{-1} Q_{ZZ} Q_{ZX}'^{-1} x_{i_b})$ ; and this asymptotic variance can be estimated consistently by  $n\tilde{\sigma}^2 x_{i_b}' (Z'X)^{-1} Z'Z (X'Z)^{-1} x_{i_b}$ , where

$$\tilde{\sigma}^2 = (\tilde{v}'\tilde{v}/n) = (y - X\tilde{\beta} - d\tilde{\gamma})'(y - X\tilde{\beta} - d\tilde{\gamma})/n$$

is the usual consistent estimator of  $\sigma^2$ .

##### **Proof**

From (4),

$$\sqrt{n}(\tilde{\gamma} - \gamma - v_{i_b}) = (d' M_1 d)^{-1} \sqrt{n}(d' M_1 v) - \sqrt{n}v_{i_b}.$$

So, from (2), the asymptotic variance of  $\sqrt{n}(\tilde{\gamma} - \gamma - v_{i_b})$  is the same as that of  $\sqrt{n}(d' M_1 v - v_{i_b})$ .

Now,

$$\begin{aligned} asy.var.[\sqrt{n}(d' M_1 v - v_{i_b})] &= asy.var.[\sqrt{n}(d' v - d' X(Z'X)^{-1} Z'v - v_{i_b})] \\ &= asy.var.[x_{i_b}' (n^{-1}Z'X)^{-1} (n^{-1/2}Z'v)] \\ &= x_{i_b}' Q_{ZX}^{-1} (\sigma^2 Q_{ZZ}) Q_{ZX}'^{-1} x_{i_b}, \end{aligned}$$

which completes the first part of the proof.

Next, consider  $\tilde{\sigma}^2 = (\tilde{\nu}'\tilde{\nu} / n)$  as an estimator of  $\sigma^2$ . Noting that  $d'd = 1$ ,

$$\begin{aligned}\tilde{\sigma}^2 &= n^{-1}\nu'\nu + n^{-1}(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta) + n^{-1}(\tilde{\gamma} - \gamma)'(\tilde{\gamma} - \gamma) + 2n^{-1}(\tilde{\gamma} - \gamma)'d'X(\tilde{\beta} - \beta) \\ &\quad - 2n^{-1}(\tilde{\gamma} - \gamma)d'\nu - 2n^{-1}(\tilde{\beta} - \beta)X'\nu.\end{aligned}\tag{10}$$

By Khintchine's Theorem,

$$plim(n^{-1}\nu'\nu) = \sigma^2.\tag{11}$$

Also, using Slutsky's Theorem repeatedly:

$$plim[n^{-1}(\tilde{\gamma} - \gamma)'d'\nu] = plim(\nu_{i_b}^2 / n) = 0;\tag{12}$$

$$plim[n^{-1}(\tilde{\beta} - \beta)'X'\nu] = 0 \quad plim(n^{-1}X'\nu) = 0;\tag{13}$$

$$plim[n^{-1}(\tilde{\gamma} - \gamma)'d'X(\tilde{\beta} - \beta)] = plim(\nu_{i_b}x_{i_b} / n)plim(\tilde{\beta} - \beta) = 0;\tag{14}$$

and

$$plim[n^{-1}(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta)] = plim[(\tilde{\beta} - \beta)'Q_{XX}(\tilde{\beta} - \beta)] = 0.\tag{15}$$

In addition, using (2):

$$\begin{aligned}plim[n^{-1}(\tilde{\gamma} - \gamma)'(\tilde{\gamma} - \gamma)] &= plim[n^{-1}\nu'M_1'dd'M_1\nu] \\ &= plim[n^{-1}\nu'dd'\nu - 2n^{-1}\nu'Z(X'Z)^{-1}X'd'd\nu \\ &\quad + n^{-1}\nu'Z(X'Z)^{-1}X'dd'X(Z'X)^{-1}Z'\nu] \\ &= plim[\nu_{i_b}^2 / n] - 2plim[n^{-1}\nu'Z]Q_{ZX}^{-1}plim[n^{-1}x_{i_b}\nu_{i_b}] \\ &\quad + plim[n^{-1}\nu'Z]Q_{ZX}^{-1}plim[x_{i_b}^2 / n]Q_{ZX}^{-1}plim[n^{-1}Z'\nu] = 0.\end{aligned}\tag{16}$$

Then, using (11) to (16) in (10),  $plim(\tilde{\sigma}^2) = \sigma^2$ .

Finally, let  $\tilde{V}[\sqrt{n}(\tilde{\gamma} - \gamma - \nu_{i_b})] = n\tilde{\sigma}^2x_{i_b}'(Z'X)^{-1}Z'Z(X'Z)^{-1}x_{i_b}$ . Immediately, from the consistency of  $\tilde{\sigma}^2$ ,

$$\begin{aligned}
plim\{\tilde{V}[\sqrt{n}(\tilde{\gamma} - \gamma - v_{i_b})]\} &= \sigma^2 x_{i_b}' (n^{-1} Z' X)^{-1} (n^{-1} Z' Z) (n^{-1} X' Z)^{-1} x_{i_b} \\
&= \sigma^2 x_{i_b}' Q_{ZX}^{-1} Q_{ZZ} Q_{ZX}^{-1} x_{i_b} \\
&= asy.var.[\sqrt{n}(\tilde{\gamma} - \gamma - v_{i_b})] .
\end{aligned}$$

■

So, in the model that we are considering, although the IV estimator for the (scalar) coefficient of the dummy variable is inconsistent, the corresponding estimator of its asymptotic variance is a consistent estimator of the true asymptotic variance of  $\tilde{\gamma}$ . In particular, note that this estimator is the one that would usually be constructed in the context of IV estimation. Not surprisingly, the usual estimator for the asymptotic covariance matrix (a.c.m.) of  $\sqrt{n}(\tilde{\beta} - \beta)$  is also weakly consistent, as we now show.

**Theorem 4**

Consider model (1) and the assumptions in section 2. In addition, let  $Z_{(i_b)}$  represent the  $Z$  matrix with the  $i_b^{\text{th}}$  row deleted, and assume that  $plim(n^{-1} Z_{(i_b)}' Z_{(i_b)}) = Q_{Z^*Z^*}$  is a finite matrix. Then the a.c.m. of  $\sqrt{n}(\tilde{\beta} - \beta)$  is  $(\sigma^2 Q_{ZX}^{-1} Q_{Z^*Z^*} Q_{ZX}^{-1})$ ; and this asymptotic variance can be estimated consistently by  $n\tilde{\sigma}^2 x_{i_b}' (Z' X)^{-1} Z_{(i_b)}' Z_{(i_b)} (X' Z)^{-1}$ , where

$$\tilde{\sigma}^2 = (\tilde{v}' \tilde{v} / n) = (y - X\tilde{\beta} - d\tilde{\gamma})'(y - X\tilde{\beta} - d\tilde{\gamma}) / n$$

is the usual consistent estimator of  $\sigma^2$ .

**Proof**

From (6),

$$\sqrt{n}(\tilde{\beta} - \beta) = (n^{-1} Z' M_2 X)^{-1} (n^{-1/2} Z' M_2 v) ,$$

and from (7),

$$plim(n^{-1} Z' M_2 X)^{-1} = Q_{ZX}^{-1} .$$

So,

$$a.c.m.[\sqrt{n}(\tilde{\beta} - \beta)] = Q_{ZX}^{-1} a.c.m.(n^{-1/2} Z' M_2 v) Q_{ZX}^{-1} .$$

Now, note that  $(n^{-1/2} Z' M_2 v) = (n^{-1/2} Z_{(i_b)}' v_{i_b})$ , and so

$$a.c.m.[\sqrt{n}(\tilde{\beta} - \beta)] = \sigma^2 Q_{ZX}^{-1} Q_{Z^*Z^*} Q_{ZX}^{-1} ,$$

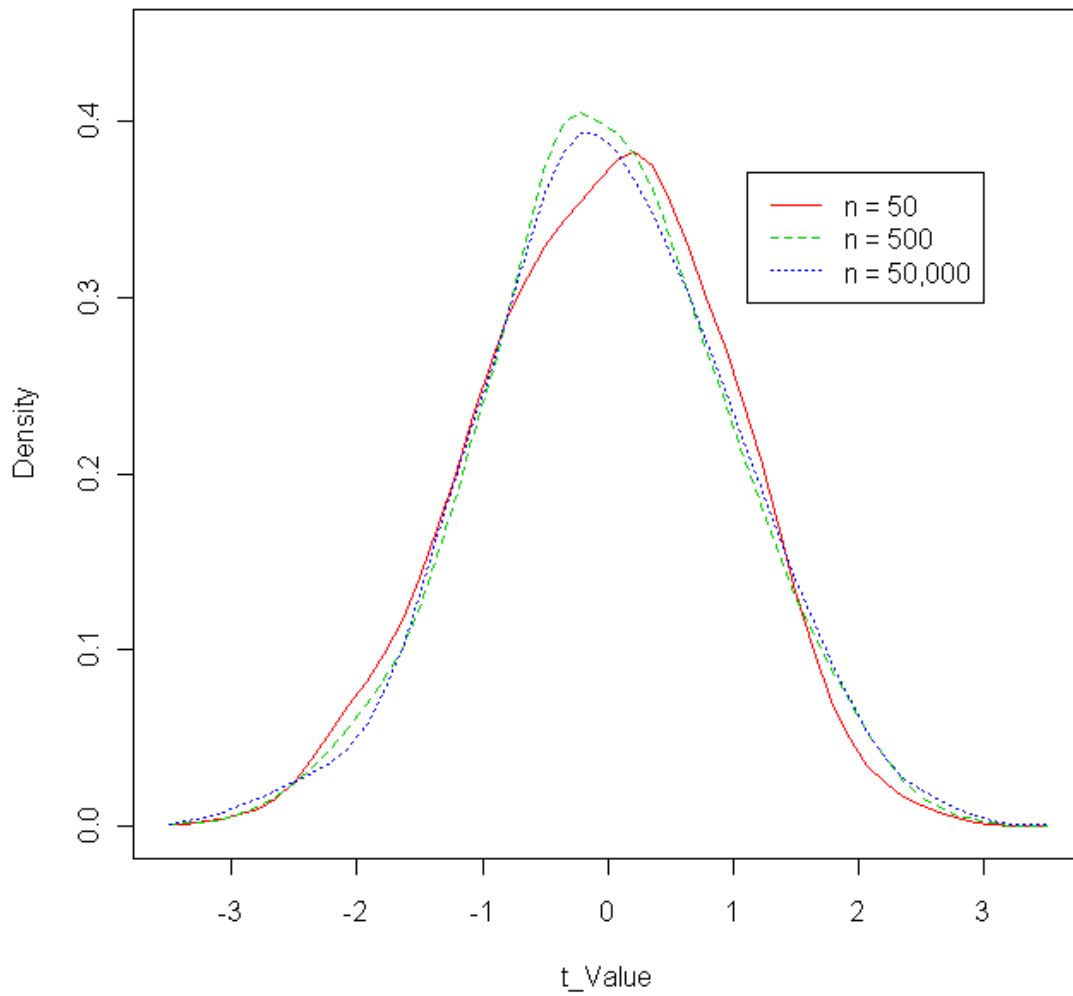
as required.

Recall from the proof of Theorem 3 that  $\tilde{\sigma}^2 \xrightarrow{p} \sigma^2$ , so immediately, by Khintchine's Theorem,

$$\begin{aligned} [n\tilde{\sigma}^2 (Z'X)^{-1} (Z_{(i_b)}' Z_{(i_b)}) (X'Z)^{-1}] &= [\tilde{\sigma}^2 (n^{-1}Z'X)^{-1} (n^{-1}Z_{(i_b)}' Z_{(i_b)}) (n^{-1}X'Z)^{-1}] \\ &\xrightarrow{p} \sigma^2 Q_{ZX}^{-1} Q_{Z^*Z^*} Q_{ZX}^{-1} = a.c.m.[\sqrt{n}(\tilde{\beta} - \beta)] \end{aligned}$$

■

**Figure 3: Sampling Distributions - Dummy Coefficient t-Statistic**



## 5. Large-sample inferences

Under our assumptions, tests of hypotheses about the elements of  $\beta$  proceed in the usual way. So, for example, the usual t-test statistic for testing  $H_0 : \beta_j = \beta_j^*$ , for the  $j^{\text{th}}$  element of  $\beta$ , is asymptotically standard normally distributed if the null hypothesis is true. Interestingly, the same is true for the corresponding t-test statistic for testing  $H_0 : \gamma = \gamma^*$ . This follows from the asymptotic normality of  $\tilde{\gamma}$  and the asymptotic independence of  $(\tilde{\gamma} - \gamma^*)$  and its estimated asymptotic variance,  $n\tilde{\sigma}^2 x_{ib}'(Z'X)^{-1}Z'Z(X'Z)^{-1}x_{ib}$ . So, asymptotically valid inferences may still be drawn about the coefficients of dummy variables that take only a fixed number of non-zero values. The usual tools apply. This is a direct (asymptotic) generalization of the Hendry and Santos (2005, p.576) result regarding the validity of the t-test in this case, in finite samples, under OLS estimation with fixed regressors and normal errors. Note that the latter condition is not needed here, and that the usual Wald test for more general exact linear restrictions on the coefficients of these dummy variables is also asymptotically valid.

Figure 3 illustrates the limiting behaviour of the t-test for testing  $H_0 : \gamma = 0$  under the set-up for the Monte Carlo experiment described in section 3, when the null hypothesis is true. The means (standard deviations) of the sampling distributions in Figure 3, in terms of increasing sample size are -0.053 (0.979), -0.027 (0.980), and -0.008 (1.007). The corresponding p-values for the Jarque-Bera test for normality are 01.020, 0.821 and 0.806.

Hendry and Santos (2005, pp.576-577) note that the usual t-test for  $H_0 : \gamma = \gamma^*$  is an inconsistent test in their framework with fixed regressors and OLS estimation. That is, for a fixed alternative and significance level, the power of the test does not approach unity as  $n \rightarrow \infty$ . This is in contrast to the usual situation with well-behaved regressors. In our case the same situation applies, essentially because the limiting variance of  $\tilde{\gamma}$  is non-zero. Table 1 illustrates this point with power curves for the asymptotic t-test of  $H_0 : \gamma = 0$ , against various fixed two-sided alternatives with a significance level of 5%. For fixed and finite  $\gamma$ , the power converges to a value less than unity as  $n \rightarrow \infty$ .

**Table 1: Powers of t-test of  $H_0 : \gamma = 0$  vs.  $H_A : \gamma \neq 0$** **(Size = 5%)**


---

$n$	$\gamma$					
	-0.5	-0.3	-0.1	0.1	0.3	0.5
50	0.37	0.19	0.06	0.05	0.12	0.31
500	0.49	0.21	0.07	0.06	0.20	0.46
5,000	0.49	0.21	0.07	0.07	0.21	0.48
100,000	0.49	0.22	0.07	0.07	0.22	0.49
200,000	0.49	0.22	0.07	0.07	0.22	0.49

---

## 6. Conclusions

The use of dummy variables which take only a single non-zero value, or whose non-zero values are fixed in number as the sample size increases, raises some interesting issues for the properties of a broad class of regression estimators. Specifically, we have shown that the members of the family of Instrumental Variables estimators will be inconsistent for the coefficients of such dummy variables. The IV estimators will, however, retain their weak consistency for the coefficients of the other (possibly random) regressors in the model, under standard assumptions. The usual estimator of the asymptotic covariance matrix for the IV estimator of the full coefficient vector is also consistent, even in the presence of such ‘special’ dummy variables. Finally, although the dummy variable coefficient estimator is inconsistent, the associated t-test statistics are asymptotically standard normally distributed, as usual. However, this test loses its usual property of consistency in the case of dummy variable coefficients. So, its power is bounded below 100% for fixed alternative hypotheses, even for infinitely large sample sizes.

Given the breadth of the family of Instrumental Variables estimators, and their widespread use in the context of both time-series and cross-sectional data, these results provide a valuable extension to the existing results in the literature. They also have immediate practical implications for applied researchers, and an example of this is provided by Giles (2011).

## Acknowledgement

I am very grateful to Martin Farnham, Ryan Godwin, and Jacob Schwartz for helpful comments and discussions.

## References

- Frisch, R. and F. V. Waugh (1933). Partial time regression as compared with individual trends. *Econometrica*, 1, 387-401.
- Giles, D. E. A. (1984). Instrumental variables regressions involving seasonal data. *Economics Letters*, 14, 339-343.
- Giles, D. E. (2011). Interpreting dummy variables in semi-logarithmic regression models: Exact distributional results. Econometrics Working Paper EWP 1101, Department of Economics, University of Victoria.
- Hendry, D. F. and C. Santos (2005), Regression models with data-based indicator variables. *Oxford Bulletin of Economics and Statistics*, 67, 571-595.
- Lovell, M. (1963). Seasonal adjustment of economic time series. *Journal of the American Statistical Association*, 58, 993-1010.
- Lovell, M. (2008). A simple proof of the FWL (Frisch, Waugh, Lovell) theorem. *Journal of Economic Education*, 39, 88-91.
- Quantitative Micro Software (2010). *EViews 7.1*, Irvine, CA: Quantitative Micro Software.
- Salkever, D. S. (1976). The use of dummy variables to compute predictions, prediction errors, and confidence intervals. *Journal of Econometrics*, 4, 393-397.