Econometrics **2015**, *3*, 55-64; doi:10.3390/econometrics3010055

OPEN ACCESS

econometrics

ISSN 2225-1146

www.mdpi.com/journal/econometrics

Article

On the Interpretation of Instrumental Variables in the Presence of Specification Errors

P.A.V.B. Swamy [1], George S. Tavlas [2] and Stephen G. Hall [2,3,*]

[1] Board of Governors of the Federal Reserve System (retired), Washington, DC, VA 22315, USA; E-Mail: swamyparavastu@hotmail.com

[2] Bank of Greece, 21 El, Venizelos Ave, 102 50 Athens, Greece; E-Mail: gtavlas@bankofgreece.gr

[3] Department of Economics, University of Leicester, Leicester, LE1 7RH, UK

* Author to whom correspondence should be addressed; E-Mail: s.g.hall@le.ac.uk; Tel.: +44-1162522827.

Academic Editor: Kerry Patterson

Received: 7 October 2014 / Accepted: 15 January 2015 / Published: 29 January 2015

Abstract: The method of instrumental variables (IV) and the generalized method of moments (GMM), and their applications to the estimation of errors-in-variables and simultaneous equations models in econometrics, require data on a sufficient number of instrumental variables that are both exogenous and relevant. We argue that, in general, such instruments (weak or strong) cannot exist.

Keywords: instrumental variables; generalized method of moments; random coefficient models

JEL classification: C11; C13

1. Introduction

Researchers are becoming increasingly aware that there are often serious problems with the use of instrumental-variable based techniques—both instrumental variable (IV) estimation and versions of generalized methods of moments (GMM) that use instrumental variables [1]. A valid instrument must be uncorrelated with the errors in an equation, that is, it must be exogeneous, and correlated with the explanatory variable, that is, it must be relevant [1,2]; [3] (p. 316); [4] (pp. 603–605). In this connection, Pratt and Schlaifer [5] pointed out that, without knowing what the errors represent, it is not possible to decide whether the exogeneity condition is correct. They also noted that the condition is "meaningless" if the errors are included in an equation to represent the net effect (on the dependent variable) of variables excluded from the equation[1]. This paper may be seen as an extension of the argument made by Pratt and Schlaifer [5] to the general case of IV estimators and, in particular, to explain why much IV estimation is plagued by either irrelevant instruments or instruments that fail the exogeneity condition. As pointed out by Murray [1] (p. 114), an instrument can be so weakly correlated with the troublesome variable that the instrument has little relevance[2].

In this paper we argue that the difficulties associated with instruments should not be surprising. Specifically, we show that valid instruments cannot exist in the presence of *any* model mis-specification. Such mis-specification can arise, indeed, is very likely to arise, from a variety of influences, including omitted variables, measurement errors, and incorrect functional forms. To generate cases in which instruments could exist, the model being estimated would have to be correctly specified; any error component of such a model would have to be a white noise process that it is independent of the instruments.

As Pratt and Schlaifer [5] make clear, the interpretation of the error in an equation is crucial here. There are two possible extreme interpretations. One interpretation is embedded in the classical regression model, which includes an error that is simply assumed to be a white noise error process with a given distribution. The alternative view is that the error is generated by all the misspecification in the model; a perfectly specified model would have no error. We would argue that the second interpretation is always more relevant in practice and it is this interpretation which gives rise to the problem with instrumental variables outlined below.

How does out framework fit with the "standard" one? The standard view typically starts from a multivariate DGP made up of a set of random variables with non-degenerate distributions. This will imply the existence of a set of error terms that are not directly associated with any misspecification in the model but which reflects the basic stochastic nature of the variables being considered. These error terms may be easily built into the analysis below simply by interpreting one (or more) of the time-varying coefficients as errors. We will not do this below, as it simply adds an extra layer of complexity without changing the results. The key assumption which makes the analysis below work, however, is that we must assume that at least part of the observed errors comes from model misspecifications, including omitted variables, measurement error and the wrong functional form. If errors are not the result of such misspecification then we would essentially be claiming to know the true model, and the criticism of instrumental variables made below will not hold true.

We would also stress that we are certainly not arguing that, in light of the problems associated with IV estimation, for a return to standard OLS, with its well-known problems. We simply show that instrumental variables do not adequately deal with these problems. There is also a reasonably large

---

1 Pratt and Schlaifer [5] go on to state that the exogeneity condition may be satisfied for certain "sufficient sets" of excluded variables. However, the point we make here is that it cannot hold for the excluded variables (in the Pratt and Schlaifer sense [5], meaning that, in principle, there are variables that should be in the equation, but are omitted; these are the excluded variables referred to by Pratt and Schlaifer [5]).

2 Additionally, it is extremely difficult to verify if an instrument is uncorrelated with the error term in the equation being estimated. For a discussion, see [6] (pp. 144–145).

literature on conducting inference in IV regressions with poor instruments; this literature includes, Cheng and Liao [7], Conley, Hansen and Rossi [8], Di Traglia [9] and Guggenberger [10]. However, this is often assuming that IV at least yields consistent estimates. We argue that this is not the case and, in general, IV is not a consistent estimator, so the accuracy of the inference made is highly questionable.

The remainder of this paper consists of three sections. Section 2 presents a general representation of model mis-specifications. We show why errors in an equation can arise. If a real-world relationship were completely known, there would be no role for a substantial error term. However, incomplete knowledge of real-world relationships is a basic component of estimated relationships. We show how correctly specified models involve time-varying coefficients (TVCs) [11], for which instruments cannot exist because, under a TVC set-up, the error terms contain the explanatory variables. Section 3 provides a simple example that illustrates our argument. Section 4 concludes.

2. A Representation of Correct Model Specification

2.1. General Considerations

In general, economic theory suggests relationships between variables, but it does not usually give clear guidance as to the correct functional form or the complete set of variables that are relevant. For example, consider an economic variable, denoted by y_t^*, and its complete set of determinants, denoted by x_{jt}^*, j = 1, …, L(t). Here the total number L(t) of determinants may be time dependent and is definitely unknown. Typically, data on y_t^* and on a subset $K - 1$ of the L(t) determinants are available. The remaining L(t) − K + 1 determinants are omitted from the model either because they are unobserved or for some other reason. Moreover, these data may contain measurement errors. Let $y_t = y_t^* + v_{0t}$ and $x_{jt} = x_{jt}^* + v_{jt}$, $j = 1, \ldots, K - 1$, where the variables without an asterisk are observable, the variables with an asterisk are unobservable true values, and vs are measurement errors. The theoretical relationship is

$$y_t^* = f_t(x_{1t}^*, \ldots, x_{L(t)t}^*) \ (t = 1, \ldots, \mathrm{T}) \tag{1}$$

with unknown functional form, no knowledge of some of the arguments of $f_t(x_{1t}^*, \ldots, x_{L(t)t}^*)$, and with no need for an error term. In other words, we do not have any omitted determinant of y_t^* in Equation (1) which is, therefore a mathematical equation. To distinguish it from a regression equation, we do not call $x_{1t}^*, \ldots, x_{L(t)t}^*$ the regressors or explanatory variables but call them the determinants of y_t^* or "the arguments" of the function $f_t(x_{1t}^*, \ldots, x_{L(t)t}^*)$. We call the arguments $x_{1t}^*, \ldots, x_{K-L(t)t}^*$ the included determinants and the arguments $x_{Kt}^*, \ldots, x_{L(t)t}^*$ omitted determinants, since data on the latter arguments are not available.

Without mis-specifying the relationship in Equation (1), we can write

$$y_t^* = \alpha_{0t} + \sum_{j=1}^{K-1} \alpha_{jt} x_{jt}^* + \sum_{g=K}^{L(t)} \alpha_{gt} x_{gt}^* \tag{2}$$

where for $\ell = j$ or g, $\alpha_{\ell t} = \dfrac{\partial y_t^*}{\partial x_{\ell t}^*}$ and $\alpha_{0t} = y_t^* - \sum_{\ell=1}^{L_t} \alpha_{\ell t} x_{\ell t}^*$, the time profiles of the $\alpha_{\ell t}$ s are determined by the correct functional form of model (1). Since the correct functional form is unknown, these time profiles are also unknown. Allowing the coefficients of Equation (2) to vary freely defines an infinite class of functional forms, which surely encompasses the correct (but unknown) functional form of Equation (2) as a special case. A main benefit of model (2) is the certainty that the infinite class of functional forms will encompass the correct functional form and, thus, the unknown functional form problem is solved.

We wish to point out that that if spline-, cubic-spline-, P-spline-, or any other-type restrictions are *imposed* on the functional form of model (1), then it can have an incorrect functional form; for examples of spline- and cubic-spline-type restrictions, see [3] (p. 111) and [12] (p. 803). A main benefit of model (2) is the certainty that the infinite class of functional forms will encompass the correct functional form. This notion, that a time varying coefficient model can exactly represent an unknown nonlinear functional form was first proved by Swamy and Mehta [13] and subsequently confirmed by Granger [14].

Clearly, the the determinants of y in Equation (2) can be correlated with each other, leading to the well-known problem of multicollinearity. In particular, the $K-1$ observable determinants (the x_{jt}^* s) in Equation (2) can be correlated with the L(t) − K + 1 omitted determinants (the x_{gt}^* s). To assume otherwise would, in the words of Pratt and Schlaifer [5], be a "meaningless" assumption. The mathematical relationship between each omitted determinant and the observed determinants is as follows

$$x_{gt}^* = \lambda_{0gt} + \sum_{j=1}^{K-1} \lambda_{jgt} x_{jt}^* \quad (g = K, \ldots, \mathrm{L(t)}) \tag{3}$$

where λ_{0gt} is a portion of x_{gt}^* remaining after the effects of the x_{jt}^* s have been removed from x_{gt}^*. Since we do not have data on the L(t) − K + 1 x_{gt}^* variables, we can eliminate them from Equation (2) by substituting Equation (3) into (2), which gives

$$y_t^* = \alpha_{0t} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{0gt} + \sum_{j=1}^{K-1} (\alpha_{jt} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jgt}) x_{jt}^* \tag{4}$$

Note that Equation (4) shows y_t^* as a function of $K-1$ included determinants and the remainders of the excluded variables, *i.e.*, what remains after subtracting the effects on the excluded variables of the $K-1$ observable determinants. Equation (4) thus solves both the unknown functional form (since it is derived from Equation (2)) and the full set of (time-varying) determinants of y_t^* in Equation (1). Thus, Equation (4) solves both the unknown functional form and omitted determinants problems. It does not, however, account for measurement errors and in this connection, we consider model (4) again, since it is not in a form that can be estimated. Such a form is derived below.

In terms of the observable variables, Equation (4) can be written as

$$y_t = \gamma_{0t} + \sum_{j=1}^{K-1} \gamma_{jt} x_{jt} \tag{5}$$

In the presence of Equation (3) and measurement errors, model (5) coincides with model (2) if

$$\gamma_{0t} = \alpha_{0t} + \sum_{g=K}^{L(t)} \alpha_{0gt}\lambda_{0gt} + v_{0t} \tag{6}$$

$$\gamma_{jt} = (\alpha_{jt} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jgt})(1 - \frac{v_{jt}}{x_{jt}}) \ (j = 1, \ldots, K-1) \tag{7}$$

According to Pratt and Schlaifer [5], the term $\sum_{g=k}^{L(t)} \alpha_{gt}\lambda_{0gt}$ in Equation (4) can be treated as an error term. With this treatment we can use the usual regression terminology from this point on.

To recapitulate, we have begun with Equation (1). To solve the unknown functional form problem, Equation (1) is replaced with Equation (2). To solve the excluded variables problem without making meaningless assumptions, Equation (3) is introduced and inserted into Equation (2) to obtain Equation (4). After introducing measurement errors at the appropriate places in Equation (4), it is replaced with Equation (5).[3] In this derivation, no approximations and no meaningless assumptions are made. The terms on the right-hand side of Equations (6) and (7) provide crucial information. Equation (4) shows that the λ_{0gt} s, in conjunction with the x^*_{jt} s, are at least sufficient to determine y^*_t. This is the proof Pratt and Schlaifer [5] (pp. 34, 50) offer to show that the second term on the right-hand side of Equation (6) is a function with the correct functional form of certain "sufficient sets" of excluded variables. The authors warn against adding an arbitrary error term to a linear or nonlinear function of the x^*_{jt} s and assuming that the x^*_{jt} s are independent of the error term.

The interpretation of the terms on the right-hand side of Equation (7) and their implications are as follows:

- The term α_{jt} is equal to $\partial y^*_t / \partial x^*_{jt}$ (if y^*_t is a continuous function of x^*_{jt}) and corresponds to the bias-free effect of x^*_{jt} on y^*_t , as can be seen from Equation (2). The right sign of α_{jt} is provided by economic theories. The correlation between y^*_t and x^*_{jt} is spurious if $\alpha_{jt} = 0$. Even though these bias-free effects are economically very meaningful, they cannot be estimated using any of the conventional econometric techniques.
- The term $\sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jg}$ captures omitted-variables bias. Note that each term in this sum is the product of two coefficients—the effect of the excluded variable x^*_{gt} on y^*_t (*i.e.*, α_{gt}) and the effect of the included variable x^*_{jt} on the excluded variable x^*_{gt} (*i.e.*, λ_{jgt}). Omitted-variable biases can exist as long as the error terms are present in econometric models.
- The term $(\alpha_{jt} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jgt})(-\frac{v_{jt}}{x_{jt}})$ captures measurement-errors bias.[4] These biases exist whenever estimates of some theoretical variables are used as explanatory variables.
- The explanatory variables of model (5) are correlated with their own coefficients because the measurement-error bias component of γ_{jt} is a function of x_{jt}.

---

[3] For the derivation, see [15].

[4] The minus sign in the expression reflects the fact that the second parenthetical term on the right-hand side of Equation (7) is one minus the ratio (v_{jt} / x_{jt}).

- Model (5) can be mis-specified if the omitted-variable and measurement-error bias (or simply, the specification bias) components of its coefficients in Equation (7) are ignored[5].

Having derived the model in Equation (5), which explicitly includes all these forms of biases, it is now possible to show why valid instruments cannot be found for this model. Combining Equations (5)–(7) into one gives

$$y_t = \alpha_{0t} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{0gt} + v_{0t} + \sum_{j=1}^{K-1}(\alpha_{jt} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jgt})(1 - \frac{v_{jt}}{x_{jt}})x_{jt} \tag{8}$$

2.2. Some Illustrative Cases

In the standard approach, we aim to choose instruments that are strongly correlated with the variable being instrumented, but which are independent of the errors in the model. If an instrument is not well-correlated with the variable under consideration, then we have the problem of weak instruments, if the instrument is not independent of the error then we will not remove the bias. We illustrate the problem with IV by considering three cases.

Case I. (Linear models). By adding and subtracting a constant parameter model we get

$$y_t = \beta_0 + \sum_{j=1}^{K-1}\beta_j x_{jt} + (\alpha_{0t} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{0gt} + v_{0t} - \beta_0 + \sum_{j=1}^{K-1}((\alpha_{jt} + \sum_{g=K}^{L(t)} \alpha_{gt}\lambda_{jgt})(1 - \frac{v_{jt}}{x_{jt}}) - \beta_j)x_{jt} \tag{9}$$

where the last two terms in Equation (9) become the error term in the model. The problem with instrumental variables in this context now becomes apparent; we need to find a variable that is both correlated with x_{jt}, but uncorrelated with the error term, which itself contains x_{jt}. Such a variable almost certainly cannot exist. We extend this proof to nonlinear models in Case III below.

Case II. (Linear errors-in-variables model without the error in equation). If

$$\lambda_{0gt} = \lambda_{jgt} = 0 \ \textit{for all } j, g \textit{ and } t \tag{10}$$

and

$$\beta_j = \alpha_{jt} \ \textit{for } j = 0, \ldots, K-1 \tag{11}$$

Equation (10) implies that there are no omitted variables and Equation (11) implies that the true model has a linear functional form. Under Equations (10) and (11), Equation (9) reduces to an errors-in-variables model and the error term becomes just $v_{0t} - \sum_{j=1}^{K-1} v_{jt}\beta_j$. For IV estimation of such a model, we need instruments that are relevant and uncorrelated with the errors (exogenous). Assumptions (10) and (11) are highly restrictive and, in effect, amount to the assumption that the model is perfectly specified and that there are no excluded variables. Hence, this extreme case rules out Pratt and Schlaifer's case [5] where the included variables are independent of the excluded variables, as there are none. The error term is then purely an identifier, in the Pratt and Schlaifer sense [5]. However we would argue that this case can never occur in the real world.

[5] Discussion of the terms in Equation (7) are provided in [16,17].

Case III. (Nonlinear models). Note that Cases I and II do not cover nonlinear models. To complete our proof of the nonexistence of valid instruments, we need to consider the (realistic) nonlinear case where model (5), with its coefficients satisfying Equations (6) and (7), holds. A natural method of identifying the coefficients of model (5) without mis-specifying its functional form is to decompose these coefficients into their respective components in Equations (6) and (7). To perform this decomposition, we assume that

$$\gamma_{jt} = \pi_{j0} + \sum_{h=1}^{p-1} \pi_{jh} z_{ht} + \varepsilon_{jt} \;\; (j = 0, 1, \ldots, K-1) \tag{12}$$

where the z_{ht} s are observable, $E(\varepsilon_{jt} \mid z_{1t}, \ldots, z_{p-1,t}) = 0$, $j = 0, 1, \ldots, K-1$, all t, and the ε_{jt} s may be serially and contemporaneously correlated. It is assumed that in model (5), the x_{jt} s are *conditionally* independent of their own coefficients given the z_{ht} s. Changes in policy variables, shift variables representing structural changes in the γ_{jt} and lagged changes in the x_{jt} s can be used as the z_{ht} s, as in [16].

We cannot be sure that the equation obtained by substituting Equation (12) into Equation (5) will have the correct functional form. The only way we can be so sure is by letting p tend to infinity so that ε_{jt} converges in probability to zero. It is possible to push ε_{jt} as low as desired with a high probability just by adding additional z_{jt} s on the right-hand side of Equation (12); it does not matter if some of the z_{jt} s are redundant in the sense that their coefficients in Equation (12) are zero. Equation (12) with infinitely large p and without ε_{jt} can explain all the variation in γ_{jt} in terms of observable variables. Substituting such an equation into Equation (5) gives an equation with the correct functional form.

Inserting Equation (12) into Equation (5) gives

$$y_t = \pi_{00} + \sum_{h=1}^{p-1} \pi_{0h} z_{ht} + \sum_{j=1}^{K-1} (\pi_{j0} + \sum_{h=1}^{p-1} \pi_{jh} z_{ht}) x_{jt} + \varepsilon_{0t} + \sum_{j=1}^{K-1} \varepsilon_{jt} x_{jt} \tag{13}$$

This is an estimable form of model (5).[6]

Now if we were to estimate a fixed coefficient IV version of Equation (5) such as $y_t = \beta_0 + \sum_{j=1}^{k-1} \beta_j x_{jt} + \omega_t$ then the error term in this equation becomes

$$\omega_t = (\pi_{00} + \sum_{h=1}^{p-1} \pi_{0h} z_{ht} + \varepsilon_{0t} - \beta_0) + (\sum_{j-1}^{K-1} (\pi_{j0} + \sum_{h=1}^{p-1} \pi_{jh} z_{ht}) + \sum_{j=1}^{K-1} \varepsilon_{jt} - \sum_{j=1}^{k-1} \beta_j) x_{jt} \tag{14}$$

The instrumental variables that are correlated with the x_{jt} s of the IV equation above, but not with the error terms of model (14), almost surely do not exist because these error terms also involve the x_{jt} s. Therefore, IV estimation is not possible.

It is sometimes claimed that lagged values of the variables in a model provide natural instrumental variables in many time-series settings. The mere fact that the value of $x_{j,t-1}$ was determined before the

6 Good approximations to the minimum variance linear unbiased estimators of the π's and the best linear unbiased predictors of the ε's can be obtained by applying an iteratively rescaled generalized least squares method to model (13). The consistency of these estimators can be established by letting T go to ∞ and letting p go to ∞ more slowly than T. For further discussion, see [15].

value of ε_{jt} should not lead one to conclude that $x_{j,t-1}$ is necessarily independent of ε_{jt}. The variable $x_{j,t-1}$ may well have been influenced by a forecast of a variable represented in ε_{jt}, or both $x_{j,t-1}$ and ε_{jt}, may have been affected by some third variable, as shown by Pratt and Schlaifer [5] (p. 47). Of course, if $x_{j,t-1}$ were independent of the error then this would imply that it was no longer relevant.

3. A Simple Example

Consider a simple example where the only misspecification is measurement error in the independent variable. Assume that we have a perfectly fitting linear relationship in the true variables:

$$y_t^* = \beta x_t^* \tag{15}$$

where the measured value of x_t is given by

$$x_t = x_t^* + v_t \tag{16}$$

then, the model we estimate is

$$y_t^* = \beta x_t - \beta v_t \tag{17}$$

where $-\beta v_t$ is an error term.

There are two ways we can demonstrate the problem with IV applied to Equation (17). First, we may consider the issue from a TVC perspective and we write an exact version of Equation (15) as

$$y_t^* = \beta_t x_t \tag{18}$$

where $\beta_t = \beta(1 - \frac{v_t}{x_t})$. Here we avoid the assumption that there exists an instrument, denoted by q, such that it is correlated with x_t and uncorrelated with v_t. Then, if we apply a fixed coefficient model to this equation, we get

$$y_t^* = \beta^{**} x_t + (\beta_t - \beta^{**}) x_t \tag{19}$$

where we are only considering the cases in which $\beta^{**} \neq \beta$ so that the last term is not the same as the last term in Equation (17). The last term in Equation (19) is the error term. We can see that no valid instruments can exist for x_t since x_t is also in the error term.

We can also show the same problem from a more conventional perspective. If we perform a fixed coefficient regression, then we can rewrite Equation (15) as

$$y_t^* = \beta^1 x_t + (\beta x_t^* - \beta^1 x_t) \tag{20}$$

where the term in brackets is the error term. We again can see that the error term contains the same variable that we are trying to instrument. Thus, almost surely no valid instrument can exist.

One standard way to construct a suitable instrument[7] would be to create the following variable

$$z_t = x_{it}^* + \varepsilon_t \tag{21}$$

---

[7] We are grateful to an anonymous referee for suggesting this example.

where ε_t is uncorrelated with v_t and x_{it}^*. Let us rewrite Equation (16) as $x_{1t} = x_{1t}^* + v_t$. In this case, $[E(z_t v_t)] = [E(x_{1t}^* + \varepsilon_t)v_t] = 0$ and the instrumental variable method can yield consistent estimator of β in Equation (17). However, note that this example requires precise knowledge of the misspecification that we are trying to correct for. In other words, if we can construct z_t then we know x_{it}^* and so the whole problem goes away as we could simply have estimated Equation (15) without any measurement error and, therefore, IV would have been unnecessary in this case.

4. Conclusions

The instrumental variables that are correlated with the x_{jt}s of model (5), but not with the error terms of model (13), do not, in general, exist because these error terms also involve the x_{jt} variables. These arguments help explain why practical work with IV methods is plagued by several problems. We would argue that a much better way forward in terms of practical estimation rests on avoiding incorrect functional forms and recognition of the potential sources of omitted-variable and measurement-error biases which are present in Equation (5). By accounting for these sources of biases, we are able to show that (i) the unknown functional form give rise to TVCs; and (ii) in this TVC set-up, instruments almost surely cannot exist.

Acknowledgment

We wish to thank the editor, Kerry Patterson, and three referees for their thoughtful comments on earlier drafts of this paper.

Author Contributions

All authors made equal contributions.

Conflicts of Interest

The authors declare no conflicts of interest

References

1. Murray, M.P. Avoiding invalid instruments and coping with weak instruments. *J. Econ. Perspect.* **2006**, *20*, 111–132.
2. Staiger, D.; Stock, H. Instrumental variables regression with weak instruments. *Econometrica* **1997**, *65*, 557–586.
3. Greene, W.H. *Econometric Analysis*, 6th ed.; Pearson/Prentice Hall: Upper Saddle River, NJ, USA, 2008.
4. Westoff, F. *An Introduction to Econometrics: A Self-Contained Approach*; MIT Press: Cambridge, MA, USA, 2013.
5. Pratt, J.W.; Schlaifer, R. On the interpretation and observation of laws. *J. Econom.* **1988**, *39*, 23–52.
6. Kennedy, P. *A Guide to Econometrics*, 6th ed.; Blackwell Publishing: Malden, MA, USA, 2008.

7. Cheng, X.; Liao, Z. Select the Valid and Relevant Moment: An Information Based LASSO for GMM with Many Moments. Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2343130 (accessed on 7 October 2014).
8. Conley, T.G.; Hansen, C.B.; Rossi, P.E. Plausibly exogenous. *Rev. Econ. Stat.* **2012**, *94*, 260–272.
9. Di Traglia, F. Using Invalid Instruments on Purpose: Focused Model Selection and Averaging for GMM. 2014, Unpublished work.
10. Guggenberger, P. On the assymptotic size distortion of tests when instruments locally violate the exogeneity assumption. *Econom. Theory* **2012**, *28*, 387–421.
11. Swamy, P.A.V.B.; Tavlas, G.S. Random coefficient models. In *A Companion to Theoretical Econometrics*; Baltagi, B.H., Ed.; Blackwell: Malden, MA, USA, 2001.
12. Judge, G.G.; Griffiths, W.E.; Hill, R.C.; Lukepohl, H.; Lee, T. *The Theory and Practice of Econometrics*, 2nd ed.; John Wiley and Sons: New York, NY, USA, 1985.
13. Swamy, P.A.V.B.; Mehta, J.S. Bayesian and non-Bayesian analysis of switching regressions and a random coefficient model. *J. Am. Stat. Assoc.* **1975**, *70*, 593–602.
14. Granger, C.W.J. Nonlinear models: Where do we go next, time varying parameter models? *Stud. Nonlinear Dyn. Econom.* **2008**, *12*, 1–9.
15. Swamy, P.A.V.B.; Tavlas, G.S. The New Keynesian Phillips Curve and inflation expectations: Re-specification and interpretations. *Econ. Theory* **2007**, *31*, 293–306.
16. Hall, S.G.; Swamy, P.A.V.B.; Tavlas, G.S. Generalized cointegration: A new concept with an application to health expenditure and health outcomes. *Empir. Econ.* **2012**, *42*, 603–618.
17. Swamy, P.A.V.B.; Hall, S.G.; Tavlas, G.S. Microproduction functions with unique coefficients and errors: A reconsideration and respesification. *Macroecon. Dyn.* 2015, Unpublished work.