Causal misspecifications in econometric models

Itkonen, Juha

ifo Institute for Economic Research, HECER

27 May 2011

Helsinki Center of Economic Research

Discussion Papers

Causal misspecifications in econometric models

Juha Itkonen
Ifo Institute for Economic Research and HECER

Discussion Paper No. 327
May 2011

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014
University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 327

Causal misspecifications in econometric models*

Abstract

We bridge together the graph-theoretic and the econometric approach for defining causality in statistical models to consider model misspecification problems. By presenting a solution to disagreements between the existing frameworks, we build a causal framework that allows us to express causal implications of econometric model specifications. This allows us to reveal possible inconsistencies in models used for policy analysis. In particular, we show how a common practice of doing policy analysis with vector error-correction models fails. As an example, we apply these concepts to discover fundamental flaws in a resent strand of literature estimating the carbon Kuznetz curve, which postulates that carbon dioxide emissions initially increase with economic growth but that the relationship is eventually reversed. Due to a causal misspecification, the compatibility between climate and development policy goals is overstated.

JEL Classification: C50, Q54, Q43

Keywords: Causality, Policy evaluations, Energy consumption, Carbon dioxide emissions, Economic growth, Environmental Kuznets curve.

Juha Itkonen

Ifo Institute for Economic Research
Poschingerstr. 5
81679 Munich
Germany

e-mail: juha.itkonen@helsinki.fi

* Many thanks to Silke Friedrich, Hannu Kivimäki, Markku Lanne, Karen Pittel, Panu Poutvaara, and Harri Turunen for helpful critique, comments, and discussion. Funding and support from HECER and Ifo Institute is gratefully acknowledged.

1. Introduction

To specify a model with causality is one of the most fundamental tasks of econometric policy analysis. Yet it seems that relatively little effort is put into developing analytical tools to help assess causality in econometric models and to gain insights on the *a priori* implications of different model specifications. The probabilistic language commonly used in econometrics does not provide a rigorous definition of causality. Therefore it is often exhaustingly difficult to debate about causes and effects in complex models with many equations, variables, and statistical assumptions.

Scientific discourse on causality requires a rigorous framework that allows to describe systematically, to break down, and to assess the meaning of various statements regarding causality between variables. Such *causal frameworks* allow to express well-defined causal statements that have well-defined truth values.

Different frameworks for handling causality have been developed (Hoover, 2008). The potential-outcome framework developed by Neyman (1935), Rubin (1974, 1978), and Holland (1986) is mostly used within statistics (Pearl, 2009; Heckman, 2005) and the program evaluation literature (Angrist and Krueger, 1999; Imbens and Wooldridge, 2009). It defines causality by an analogy with random experiments. Cowles Commission econometricians (Haavelmo, 1943, 1944; Marschak, 1950; Simon, 1953) defined causality with respect to structural equations models. The elaborate graph-theoretic framework by Judea Pearl (2009), which has received much attention recently, is partly based on the definition of causality by the Cowles group. The "econometric approach" for defining causality by Heckman (2005, 2008, 2010), and Heckman and Vytlacil (2007), which also builds on the Cowles Commission econometrics, has one crucial advantage over Pearl (2009) and the potential-outcome framework: it allows for simultaneous determination of variables (nonrecursive models and cyclic graphs) which is a property often needed in econometrics. The potential-outcome framework and Cowles Commission econometrics are considered to be special cases of Pearl's framework (Pearl, 2009) and the econometric approach (Heckman and Vytlacil, 2007). Note that these interpretations of causality have little to do with predictive relationship coined as Granger causality (e.g. Granger, 1969).

A main shortcoming of potential-outcome framework and the econometric approach is their inability to express multi-level causal orderings between variables and represent them as graphs, as done by the Cowles group and

Pearl (2009). Causal ordering implied by a econometric model can be critical for the interpretation of the model. Visualizing causal relationships using graphs is not merely a way to illustrate a model, but also a tool for rigorous deduction.

Pearl's framework has received some criticism by econometricians (Leroy, 2001; Neuberg, 2003; Hoover, 2003; Heckman, 2005; Heckman and Vytlacil, 2007; Heckman, 2008), with some resolved (Pearl, 2003; Pearl, 2009, p. 374–380) and some left open. Its remaining weakness is the need to attach an unique dependent variable to each equation to define causality, which is seldom justified by economic theory.[1]

To get the best parts of both the econometric approach and Pearl's framework, we build on a theory by Herbert Simon (1953) to establish a language to describe causal relationships and extract them from econometric models. We define a causal framework that fixes the aforementioned weaknesses in previous frameworks.

Our framework allows us to define and detect causal misspecification problems in econometric models. That is, it enables us to discover inconsistencies within a wide range of models that are used for policy analysis. With these tools we can tackle problems in more complex models where they have been left unnoticed. We use the tools to show how policy analysis with vector error-correction models (VECM) (Engle and Granger, 1987) gives inconsistent and biased results under common practices (see e.g. Lütkepohl, 2005; Watson, 1994; Hamilton, 1994; Johansen, 1995). The problem is related to the asymmetry of the cointegration rank tests.

To apply the tools into practice, we analyze a resent strand of literature estimating the carbon Kuznetz curve (CKC), which postulates an "inverted U"-shape relation between carbon dioxide emissions and aggregate output of economies. That is, emissions are assumed to initially grow with output, but the relation is reversed when the economy reaches a higher level of output. In this paper we discuss the problems related to the new strand of literature by focusing on the seminal work by Ang (2007). We argue, first, that the model includes an implicit constraint that is neglected in the estimation. Second, the causal effects are misinterpreted due to variable definitions. Third, the

[1]For example, with a market equilibrium one would have to argue that one equation (demand) determines the price and another equation (supply) determines the quantity. On the contrary, most economists would consider that price and quantity are determined simultaneously by demand and supply.

model is biased to overstate the compatibility of development and environmental policy goals. In a more realistic model emissions rise quicker, peak later, and decrease slower as output increases.

It is our intention to strictly limit to the task of analyzing model specifications and their consistency.[2] We do *not* discuss identification, estimation, or testing, except to point out immediate implications regarding them, and to show the relevance of the theoretical considerations. Forgetting this distinction is a source of much confusion (Heckman and Vytlacil, 2007; Pearl, 2009).

In the next section we present a framework for expressing causality. In the third section we introduce the new CKC literature, discuss Ang's (2007) model, and derive an accounting identity that causes one of the problems. In the fourth section we describe the aforementioned three problems and apply the causal framework. In the fifth section we conclude.

2. Causal framework

Next we develop a framework for analyzing causality. We present the parts of the causal framework that are needed to show how common practice policy analysis with VECMs fails and to present the misspecification problems in the new CKC literature. We build on the graph-theoretic, causal Bayesian network literature pioneered by Pearl (2009), but we take seriously the criticism made by econometricians (Leroy, 2001; Neuberg, 2003; Hoover, 2003; Heckman, 2005; Heckman and Vytlacil, 2007; Heckman, 2008) and adapt ideas suggested by Heckman and Vytlacil (2007). To define causality, we utilize and generalize the theory by Simon (1953) (which is also augmented by Simon and Rescher (1966); Simon and Iwasaki (1988); Druzdzel and Simon (1993)).

We define a causal framework that overcomes weaknesses in Pearl's causal model and Heckman's "econometric approach", in order to enhance its suitability for econometric analysis. The main benefit is that we can analyze causality with graphs in models with simultaneously determined variables, such as the vector autocorrelation (VAR) model. Next we give a descriptive account of the distinctive features of our framework before giving a rigorous definition.

[2] This is distinguished as the task of "defining counterfactuals" in Heckman (2005, 2008, 2010), and Heckman and Vytlacil (2007).

First, in order to analyze causality as defined in the Bayesian network literature, Pearl's (2009) framework is limited to models with acyclic relationships, or recursive systems, which means there are no cycles of causation. For such models, defining causal interpretations for parameters is less debated (Strotz and Wold, 1960, 1963; Basmann, 1963b,a). However, most econometric models are nonrecursive, i.e. some variables are determined simultaneously by a set of equations, and there can be causal cycles. When variables are simultaneously determined, Pearl's causal model can not be used. Pearl's limitation to recursive systems has been the main point of criticism by econometricians (Leroy, 2001; Hoover, 2003; Neuberg, 2003; Heckman and Vytlacil, 2007).

We, however, are interested in the *necessary* causal relationships implied by a econometric model, as we are looking for inconsistencies in model specifications. Therefore we leave open the question of causality between simultaneously determined variables and focus on what can be agreed upon. This allows us to extend the framework to nonrecursive models.

Second, nonrecursiveness of models implies that variables enter model equations symmetrically. This means we circumvent the so called *matching problem*, i.e. the need to identify one variable for each equation and consider that variable as the output (or the dependent variable) of the equation.[3] Such matching has been a major obstacle for using Pearl's framework in econometrics, since economic theory usually does not justify that there is a specific dependent variable for each equation of the model.[4]

Our way of avoiding the matching problem, however, makes it impossible to use Pearl's definitions for interventions and counterfactuals. In Pearl's framework interventions and counterfactuals are defined by "surgery" and "shutting down" equations (Pearl, 2009, p. 374). Anyhow, such model manipulations have received criticism from econometricians (Neuberg, 2003; Heckman and Vytlacil, 2007).

Third, in our framework interventions and counterfactuals are defined with variables that are determined outside the models (see Pearl, 1993, 2009, p. 70–72). The points of interventions are modelled explicitly and separately by defining a set of possible *actions*. We consider models as isolations and

[3]Instead, we partition the sets of equations (derived substructures, defined later) and variables (which are solved in the derived substructures), and match between these subsets (see Dash and Druzdzel, 2008).

[4]For example, this is the case with market equilibria.

idealizations of the world (Mäki, 2011, 2008) and regard the economist as a deistic demiurge of the model world: first the economist defines the model that represents reality, next he chooses an action (a policy or a treatment) outside of the model, and then lets the model resolve itself generating an outcome. Consequently we manage without Pearl's *do*-operator and "surgery". We note that to choose nothing is also an action, and hence interventions can be defined equivalently using conditional probabilities or fixing (Pearl 2009, p. 70–72, Heckman and Vytlacil, 2007).

The action variables are strictly separated from other variables. Hence, the demiurge can never fix the same variables that describe the unit of observation. For example, the price determined by the market is never quite the same variable as the price determined by the government. However, if justified, the model can be specified such that fixing an action to a particular value (government-set price) has the same result as if other variables (market price) realize with the same value. This strict separation of action variables and other variables is pivotal for defining interventions without model manipulations like the "surgery".

The econometric approach ends up doing surgery in a very implicit way, as argued by Pearl (2009, p. 374–380). One simply can not fix endogenous variables (to assess causal effects) when their value is uniquely determined by the set of model equations: If one changes the value of the outcome to something other than the solution, it results in a contradiction with the set of equations. By definition there is only one outcome that satisfies the set of equations. This somewhat trivial notion is often neglected in the literature. To assess the causal effect between endogenous variables, one needs to explicitly formulate and justify a new, consistent causal model, which isolates and idealizes reality in a way that the causing variable is an action variable.[5] In our framework "surgery" *can* be expressed by specifying the points of intervention and action variables accordingly. That is, equations can be shut down with action variables if so specified. Moreover, we never end up contradicting with the initial model. The resulting framework is a slight generalization of Pearl (2009, p. 71).

Fourth, assuming symmetry between variables makes extracting a causal ordering slightly more complicated compared to acyclic models. We develop further the theory introduced by Simon (1953), known as the *causal ordering*

[5]A similar notion is made by Strotz and Wold (1960).

algorithm, which defines a systematic way to express the asymmetric relationships between variables in a set of equations, relationships that correspond to the conventional meaning of causality.[6] That is, a set of equations related to a econometric model contains in its formulation a description of causal relationships, and Simon's theory allows us to explicate this causal content. We generalize Simon's (1953) theory into a non-linear and nonparametric form, we simplify it by adopting a new conceptualization, and proof that the causal ordering is unique.

Fifth, the main weakness of the econometric approach is its limitation to a simple dichotomic interpretation of causal ordering: variables are either effects or causes. Using causal graphs we can form a more complex picture where variables can be both causes of some variables and effects of others. Such a mesh of relationships can be easily depicted by a directed graph. This also obviates the need to differentiate between ex ante and ex post, or objective and subjective outcomes as done by Heckman and Vytlacil (2007); Heckman (2010).[7]

Our concept of "action variable" is similar to "treatment" in Heckman and Vytlacil (2007), except that a treatment can depend on other variables (for example in "causal models for the choice of treatment"). In our framework such a model would be expressed by considering the treatment as an endogenous variable.

Sixth, because we only consider model specifications and set aside the other tasks of statistical inference, we can avoid a bountiful source of confusion by considering models simply as well-defined mathematical objects. Hence problems with *policy invariance* (as defined e.g. by Leamer, 1985; Heckman and Vytlacil, 2007), *autonomy* (as defined e.g. by Haavelmo, 1944; Aldrich, 1989; Pearl, 2009), and endogeneity (as defined e.g. by Wooldridge, 2002, p. 50–51) are not attributes of a single model, but are defined as a *relationships* between two models. When considering bias, we need to spec-

[6]Note that in this framework causality is defined in reference to a particular model specification, and does not necessitate causality in the real world nor include all forms of causality in theory: First, the framework itself is not subject to Hume's critique, problems of observing causality, or any ontological problems (Simon, 1953). Second, determination of variables by definition is also defined as causal. Third, causal relationships between simultaneously determined variables are not determined. Here causality is simply a concept that proofs very useful in analyzing problems with model specifications.

[7]We do not explicate ex ante and ex post outcomes as they simply are special cases.

ify the “true” and the “biased” model.[8] A model is called “biased”, if there exists a better justified model, the “true” model, in which a given parameter of interest (e.g. the causal effect) has a different value. While postulating the “true” model, one also confirms that the offered specification is without contradictions.

Moreover individual models are always policy invariant and autonomous. That is, the *parameters* of a model do not vary because they are fixed by definition. If one wants to consider variation in these values, one needs to specify a model where they are defined as *variables*. Similarly models (as such) never suffer from endogeneity problems because all dependences between variables are to be specified by the model. This does not, however, imply that we only analyze “structural” models (in any sense of the word surveyed by Heckman and Vytlacil (2007)).

Furthermore, a causal parameter of interest might be *infeasible* in the “true” model, i.e. the parameter can not be defined in the “true” model without contradictions, even if it is well-defined in the “biased” model.[9] In such cases the “biased” model does not just give distorted estimates, but the very existence of such an estimate is in logical conflict with the “true” model. This problem is far more severe than a simple case of poor approximation.

A weakness of our framework is that it is designed only for analyzing model specifications, so it makes no accommodations for identification, estimating, or testing. However it is clear that these tasks fundamentally depend on the soundness of the model specification.

We proceed as follows: first, we define a causal model which states the necessary specifications to consider causality in an econometric model, second, we show how to derive causal ordering between variables in a causal model, third, we show how to depict the causal ordering as a causal graph, fourth, we give an example, fifth we define causal effects and causal misspecification bias, and finally we consider the case of VECMs.

[8]The idea is similar to the distinction between small and middle-size worlds by Simon and Iwasaki (1988). See also Mäki (2011).

[9]For example, suppose that the causal effect is calculated between variables x and y in the “biased” model. If x and y are both endogenous in the “true” model, then their value is determined by the model and can not be changed by the demiurge. Hence the causal effect can not be defined.

2.1. Causal model

Let us consider a *statistical model* that consists of a set of random variables W and a set of probability distributions P satisfying *model equations*

$$g_i(w_1, \dots, w_n) = 0, \quad i = 1, \dots, m, \tag{1}$$

where g_i's are real-valued functions; $w_j \in W$ for all $j = 1, \dots, n$; and $n \geq m$.

Suppose we can partition the set W of a statistical model into disjoint subsets Y, X and U, where Y has m members. This could be called a econometric model (see e.g. Matzkin, 2007). Variables Y are observed and determined inside the model, and named endogenous; variables X are observed and determined outside the model, and are usually named exogenous; and variables U are unobserved and determined outside the model (as in Leamer, 1985). Here, we do not differentiate between unobserved and exogenous variables, because they have identical roles when identification is not considered (Heckman and Vytlacil, 2007), hence we include unobserved variables in set X.[10]

The set of equations (1) can depict a structural model, a quasi-structural model, a VAR model, or something completely different. We consider these equations purely mathematical entities where the equality relation is symmetric. We do not implicitly differentiate between left and right hand variables. This results in a desired ambiguity between structural and non-structural models.

Note, that some variables entering functions g_i can be trivial, i.e. not all functions depend on all variables. This independence, as seen later, is actually essential for causality.

To consider the corresponding causal model, we first specify the points of intervention for the model. Interventions are done through a set of *action variables* A, which specifies the ways the demiurge affects the model (actually or hypothetically). We define vector a^* as a particular value of A which represents the action of making no intervention. We express the points of intervention with *deep* model functions f_i, which correspond to model (1) when $a = a^*$, i.e.

$$f_i(w, a^*) = g_i(w)$$

[10]Pearl (2009) partitions the variables into endogenous variables and background variables, which are similar to exogenous and unobserved variables, but are used in a slightly different role due to a different definition of model equations.

is satisfied for all $w = (w_1, \ldots, w_n)$ and i.[11] As a special case, we say interventions are *additive* if $f_i(w, a) = f_i(y, x, a) = g_i(y, x + a)$, where $y \in \mathbb{R}^m$, $x, a \in \mathbb{R}^{n-m}$, and $a^* = 0$. Under additive interventions, actions can be interpreted as changing the values of exogenous variables.

Next we define our fundamental concepts:

Definition 1. A set of equations S is a *structure* if it satisfies the *regularity conditions*:

1. S is uniquely solvable,
2. S contains an equal amount of equations and variables, and each subset of S contains at most as many equations as non-trivial, unsolved variables, and
3. S's subsets are uniquely solvable if and only if they have an equal amount of equations and non-trivial, unsolved variables.

Let $\mathcal{X}$ and $\mathcal{A}$ denote the sets of possible values for the sets of variables X and A, respectively.

Definition 2. A *causal model* is a 5-tuple $M = (Y, X, A, F, P)$, where Y, X, and A are sets of variables, F is a set of real-valued functions $\{f_1, \ldots, f_m\}$, and P is a set of probability functions for variables X, if for all $p \in P$, $x_0 \in \mathcal{X}$, and $a_0 \in \mathcal{A}$ the set of equations

$$f_i(y, x, a) = 0, \quad i = 1, \ldots, m \tag{2a}$$

$$x = x_0, \tag{2b}$$

$$a = a_0, \tag{2c}$$

is a structure when x_0 is in the support of p.

Compared to the plain statistical model, a causal model separates between endogenous and exogenous variables, and specifies how interventions are done via action variables. Later we will see, that this provides the required information to give meaning to causal statements.

The *exogeneity equations* (2b) and (2c) supplement the *model equations* (2a) to state which variables are determined outside of the model. This is

[11] Similar definitions are given by Heckman (2005) when naming the set of actions as possible treatments and functions f_i as "deep structural" versions of g_i, and by (Pearl, 2009, p. 70–72) when defining interventions as variables.

required for an unequivocal causal ordering (Simon and Iwasaki, 1988) as shown later. Model equations (2a) and exogeneity equations (2b) and (2c) can be referred to as *mechanisms*. We emphasize that mechanisms do not necessarily relate to real world phenomena. Here a mechanism is just a name for equations of certain form.

Let vector $v(x,a) \in \mathbb{R}^n$ denote the solution (or the *outcome*) of structure (2) for given exogenous variables $x \in \mathcal{X}$ and action variables $a \in \mathcal{A}$ in causal model M. We define interventions and counterfactuals in deterministic terms:

Definition 3. Given that x occurred, changing action from a to a' has the effect of changing the outcome from $v(x,a)$ to $v(x,a')$.

Definition 4. Given that a was chosen and x occurred, the *counterfactual statement* "the outcome would have been v' if action $a' \neq a$ would have been taken" is equivalent to saying $v' = v(x,a')$.

Counterfactual statements between arbitrary subsets of variables, i.e. "outcome of variables $V_1 \subset Y \cup X \cup A$ would have been v_1, if $V_2 \subset Y \cup X \cup A$ had been v_2 (given x)" are not well defined.[12] To make such a statements, one has to add, how outcome v_2 was achieved, i.e. define which action would v_2 as an outcome. In other words, one must specify the action a for which outcome $v(x,a)$ is consistent with v_2. Therefore the truth value of such sentences depends on how the counterfactual condition was reached. If no action results in outcome v_2, such a sentence has no meaning in the given causal model.

We give an example to show how the concepts can be used. Consider a simple model equation $g(y,x,u) = y - \beta x - u = 0$, with endogenous variable $Y = \{y\}$, exogenous variables $X = \{x,u\}$, and parameter β. Assuming additive interventions, we get a deep model function $f(y,x,u,a) = y - \beta(x + a) - u$. Now we get a causal model with a structure consisting of equation

$$y - \beta(x+a) - u = 0, \quad x = x_0, \quad u = u_0, \quad \text{and} \quad a = a_0,$$

when (x_0, u_0) occurs and a_0 is chosen. The resulting outcome of the model is a vector $v(x_0,u_0,a_0) = (\beta(x_0 + a_0) + u_0, x_0, u_0, a_0)$.

[12]Similarly Heckman and Vytlacil (2007) discuss causal effects between endogenous variables.

2.2. Causal ordering

The causal ordering of a causal model can be determined by studying the structure (2). The concept of structure is a convenient way to check whether sufficient and non-contradiction information is given to consider causation in an econometric model. The key innovation of Simon (1953) is that the order of solving the structure determines the causal ordering in the model. Roughly speaking, a variable is caused by all the variables that need to be solved to determine its value. Next we develop concepts to formalize this.

Definition 5. A subset of a structure is a *substructure* if it is solvable and does not contain any proper subsets that are solvable.

The following lemma states a property of substructures, that is important for showing that the causal ordering in a structure is unique.

Lemma 1. *Substructures of a structure are disjoint.*

The proof is presented in Appendix A.

Let S be a structure and s the variables it contains. We can decompose S into two disjoint parts: First, we have S^0 which is the union of all substructures, e.g. $S^0 = S_1 \cup S_2 \cup \cdots \cup S_k$, where S_i is a substructure of structure S. Second, we have the remainder $R = S \backslash S^0$.

Definition 6. Structure S is *causally ordered* if R is nonempty.

Definition 7. Let structure $S = S^0 \cup R$ be causally ordered and let s^0 be the set of variables in S^0. Now we can solve equations S^0 to get unique values for variables in s^0. By substituting the solved variables into set R we get the *derived structure of first order* S^1.

In less formal terms, the derived structure equals set R where solved variables s^0 have been plugged in.

Under the first regularity condition, a derived structure must also be a structure. If it also is of causally ordered, its derived structure is called a derived structure of second order (of the structure S). By repeating, we can obtain the derived structure of kth order.

We proceed in this manner until the derived structure is not causally ordered, i.e. the remainder R is empty.

Definition 8. Let S^k be the derived structure of kth order structure S. The substructures of S^k are called the *derived substructures* of kth order.

The substructures of the original structure are called the derived substructures of 0th order, and their union S^0 is called the derived structure of 0th order.

Now structure S has been partitioned into disjoint subsets. We can denote

$$
\begin{aligned}
S &= S^0 \cup S^1 = (S_1^0 \cup S_2^0 \cup \cdots \cup S_{m^0}^0) \cup S^1 \\
&= (S_1^0 \cup S_2^0 \cup \cdots \cup S_{m^0}^0) \cup (S_1^1 \cup \cdots \cup S_{m^1}^1) \cup S^2 \\
&= (S_1^0 \cup S_2^0 \cup \cdots \cup S_{m^0}^0) \cup (S_1^1 \cup \cdots \cup S_{m^1}^1) \cup \cdots \cup (S_1^n \cup \cdots \cup S_{m^n}^n),
\end{aligned}
$$

where S^k is the derived structures of kth order, $S_1^k, S_2^k, \ldots, S_{m^k}^k$ are the derived substructures of kth order, m^k is the number of derived substructures of kth order, and n is the highest order of derived structures. More consisely we can write $S = \cup_{k=0}^{l} \cup_{i=1}^{m^k} S_i^k \cup S^{l+1} = \cup_{k=0}^{n} \cup_{i=1}^{m^k} S_i^k$ for any $0 < l < n-1$.

Lemma 2. *The partition is unique.*

The proof is presented in Appendix A.

Similarly the set of variables s in structure S has been partitioned according to the derived substructures in which they are solved. Due to regularity condition 3, each derived substructure with m equations must solve m variables.

As variables are uniquely related to the derived substructures, we can use the ordering to define key concepts related to variables.

Definition 9. Let variables x and y appear non-trivially in a derived substructure of kth order of structure S. Let d_x and d_y be the lowest order of derived substructure where variables x and y appear non-trivially, respectively. Then in structure S,

- if $d_x = d_y$, variables x and y are *simultaneously determined*, and
- if $d_x < d_y = k$, we say x is a *direct cause* of y.

We denote a direct causal relationship with the symbol "$\rightarrow$". Note that, the direct cause relationship is not yet transitive. Therefore we define the following:

Definition 10. Let s be the set of variables in a structure. Variable $x \in s$ is a *cause* of variable $y \in s$ if there is a sequence of variables $z_1, \ldots, z_n \in s$, for which $x \rightarrow z_1$, $z_1 \rightarrow z_2$, ... , and $z_n \rightarrow y$.

We define *causal ordering* of a structure as the set of all direct causal relationships between variables. With these definitions, as a corollary of Lemma 2, we get our main result:

Proposition 1. *The causal ordering of a structure is unique.*

This guarantees that the properties of the causal ordering relate to the structure unambiguously.

2.3. Causal graphs

Next we define a causal graph, which describes the causal ordering and which we can present visually.[13] A *causal graph* consists of a set of *nodes*, which is the set of variables a in structure A, and a set of *edges*, which is the set of all variable pairs (x, y) for which $x \to y$. Together these form a directed graph, where arrows represent direct causes. The graph is drawn such that the 0th order variables are in a column on the left, the 1th order variables are in the second column, and so on. To simplify graphs, we depict simultaneously determined variables by a left bracket, and unite all arrows leading to variables in these brackets (see Simon and Rescher, 1966).[14]

Now from Proposition 1 it immediately follows that a graph is a sufficient representation of the causal ordering of the structure:

Proposition 2. *The causal graph of a structure is unique.*

[13]This is specified as a partial causal graphs in Dash and Druzdzel (2008).

[14] In practice a causal graph of structure A can be drawn by following a simple procedure: (1) Make sure A satisfies the regularity conditions. (2) Find the substructures of A, i.e. all subsets that are solvable but themselves do not contain solvable subsets. (3) Solve the substructures. Set the solved variables in a column. If a substructure contains more than one variable, join the simultaneously determined variables by drawing a parenthesis. (4) If there are unsolved equations, plug the previously solved values in to the remaining equations. (5) Find the substructures of the set of remaining equation and solve them. Set the variables in the next column and join simultaneously determined variables by a parenthesis. (6) For each solved substructure draw an arrow from all variables that are included in the substructure but solved earlier. (7) Repeat steps 4–6 until all equations are solved.

2.4. Example

Lets look at an example to illustrate the framework. Let

$$f_1(v_1, v_3, v_5) = 0 \tag{3a}$$
$$f_2(v_2, v_3, v_4) = 0 \tag{3b}$$
$$f_3(v_3, v_4, v_5) = 0 \tag{3c}$$
$$f_4(v_3, v_4, v_6) = 0 \tag{3d}$$
$$v_5 = a \tag{3e}$$
$$v_6 = b \tag{3f}$$

be a structure, where a and b are given. First, we see that equations (3e) and (3f) are substructures and form the derived substructures of 0th order. Therefore the structure is causally ordered. Variables v_5 and v_6, which could be exogenous or action variables, can be solved and substituted in the other equations. Next, while v_5 and v_6 are now given, equations (3c) and (3d) can be solved simultaneously, but not separately. This means $\{(6b), (6c)\}$ is the only derived substructure of 1th order. Therefore values for v_3 and v_4 are simultaneously determined. Finally, we are left with equations (3a) and (3b). They both form a derived substructures of 2th order. Solving them gives the value for v_1 and v_2.

With this information we can form a graph:

$$\begin{tikzcd} v_5 \arrow[rr, bend left] \arrow[dr] & & v_1 \\ & \left\{\begin{matrix} v_3 \\ v_4 \end{matrix}\right. \arrow[ur] \arrow[dr] & \\ v_6 \arrow[ur] & & v_2. \end{tikzcd} \tag{4}$$

In graph (4) we see, for example, that v_6 is a cause of v_1 and a direct cause of v_4. The variables v_5 and v_6 are exogenous in the structure, and variables v_1, v_2, v_3, and v_4 are endogenous in the structure.

2.5. Causal effects and misspecification bias

Causal misspecification is a relationship between two causal models, one of which we call the "true" and other the "biased" model. We define two types of causal misspecification: infeasibility and bias. Infeasibility occurs if the causal effect of a variable is analyzed in the "biased" model when the variable is endogenous in the "true" model. In other words, the variable is determined by the equations of the "true" model, so intervening with the variable in the "biased" model contradicts the "true" model (see an example below). Bias occurs when the magnitude of causal effect is different in the "true" and "biased" models.

To be more specific, we define the individual level deterministic causal effect of a change in an action for a causal model with continuous variables and differentiable model functions. The *causal effect (CE)* of action $a \in A$ on $v \in Y \cup X \cup A$ of a causal model is the total derivative $\frac{\mathrm{d}v}{\mathrm{d}a}$. It can be calculated from the structure of the causal model by using the implicit function theorem (or taking the total derivatives and applying the Cramer's rule) (see Appendix B).

With additive interventions, an action a enters the functions similarly to it's corresponding exogenous variable $x \in X$. Therefore the derivative is same for both the action and exogenous variable. That is, for any variable $v \in Y$ applies $\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{\mathrm{d}v}{\mathrm{d}a}$. This means we can analyze a model as if the exogenous variables could be affected directly, which simplifies the causal model significantly. (An example will be given in Appendix F.)

For the rest of the paper we specify all models with additive interventions. Therefore we can analyze models as if the demiurge chooses the exogenous variables directly and omit the action variables from the structures and the causal graphs.

Let $M_1 = (Y_1, X_1, A_1, F_1, P_1)$ be a "true" causal model with continuous variables and differentiable model functions with additive interventions.[15] Now consider a similar "biased" causal model $M_2 = (Y_2, X_2, A_2, F_2, P_2)$ such that there exists variables $x \in X_1, X_2$ and $y \in Y_1, Y_2$. Suppose CE_1 and CE_2 are the causal effect of x on y in models M_1 and M_2, respectively. Then the *causal misspecification bias* of the causal effect is $CE_2 - CE_1$.

For example, such a bias emerges when the causal graph of the biased model does not include all causal paths from the effect x to the cause y. This occurs when the "biased" model incorrectly assumes a variable, z, as exogenous when it actually is endogenous such that x is a cause of z and z is a cause of y. These missing paths, i.e. chains of arrows from x through z to y, are easy to detect from the causal graphs, as is with case of VECM.

[15] The "true" model parallels to the "all causes" model in Heckman and Vytlacil (2007).

2.6. Causal misspecification bias in VECM

Consider a vector autoregression model with three variables y_t^1, y_t^2, and y_t^3, one lag, and a error-correction representation

$$\Delta y_t^i = b_i^1 z_{t-1}^1 + b_i^2 z_{t-1}^2 + \varepsilon_t^i, \tag{5}$$

where ε_t^i are random disturbances;

$$\Delta y_t^i \equiv y_t^i - y_{t-1}^i \tag{6}$$

are differences;

$$z_t^1 \equiv a_1^1 y_t^1 + a_1^2 y_t^2 + a_1^3 y_t^3 \tag{7}$$
$$z_t^2 \equiv a_2^1 y_t^1 + a_2^2 y_t^2 + a_2^3 y_t^3 \tag{8}$$

are error-correction terms for every time $t = 0, \ldots, T$; and b_i^j and a_j^i are parameters for all $i = 1, 2, 3$ and $j = 1, 2$. (See for example Hamilton, 1994, p. 580–582.) We call this model B.

Now consider a related model A, where equation (8) is substituted with $z_t^2 = 0$. This corresponds to omitting the second cointegration equation from the model.

Finally, consider model C, which is similar to model B in equations (5)–(8) but includes a third cointegration equation. Hence, equations (5) are supplemented with a third error error-correction term z_t^3.

We can summarize the three models by noting that they are all VECMs where the cointegration rank of model A, B, and C are 1, 2, and 3, respectively. That is, the models have a different number of cointegration equations included in them.

A common practice of doing policy analysis with VECMs is to consider a part of the model, called the long-run model (see e.g. Engle et al., 1989), which consists of equations (7)–(8) in the case of model B, i.e. the cointegration equations. When considering the long-run model, the error-correction terms, z_t^i, are considered as exogenous error terms. Policy assessment is made by considering counterfactual statements with respect to these equations and evaluating causal effects between the variables.

To make such policy analysis, one must first decide which cointegration equations to include in the model specification, i.e. one needs to specify the cointegration rank (or choose on of models A, B, and C). Different sets of cointegration equation give different values for causal effects.

The problem with the common practice is in the statistical test procedure (see e.g. Watson, 1994; Hamilton, 1994; Lütkepohl, 2005) based on Johansen (1988, 1995), to specify the cointegration rank. More specifically, the problem is with the tests for cointegration rank that are used to determine which cointegration equations (7)–(8) to include in the model. In the procedure, first, we justify the existence of one cointegration equation by testing the null hypothesis of no cointegration equations against the null hypothesis of one or more cointegration equations. If the null hypothesis is rejected, we infer that there must be at least one cointegration equation, in which case, we continue by testing the null hypothesis of one cointegration equation against the hypothesis of two or more cointegration equations. We continue in this way until the null hypothesis can not be rejected. Then policy evaluation is done with a model where all thus inferred cointegration equations are added.

The test is more cautious not to make a false positive, i.e not to specify the model with cointegration equations that are not in the "true" model. Correspondingly, the test makes a false negative more often, i.e. fails to reject the null hypothesis when it is false. A false negative occurs especially often when the power of the cointegration rank test is poor, like when the sample size is small and deterministic trends are possible (Demetrescu et al., 2009). This unevenness accumulates so that the test procedure tends to result in a model with a too low cointegration rank.

However when considering causal effects, causal misspecification bias occurs symmetrically with a false positive and a false negative. That is, it is equally harmful to have too many or too few cointegration equations in the model. Hence the uneven test procedure is not well-founded.

We show this symmetry by considering models A, B, and C. First, let us consider the case of a false negative. Suppose that model B is the "true" model, i.e. the true cointegration rank is 2. Suppose that in the test procedure the first null hypothesis is rejected, but the second is not. Now we infer that equation (7) holds, but equation (8) might or might not hold, and drop it out of the model. This false negative results in model A, which is "biased".

Let CE_A and CE_B be causal effects (for example between y^1 and y^2) in models A and B, respectively. Now if we use CE_A, instead of the true causal effect CE_B, we end up with causal misspecification bias $CE_A - CE_B$ because of the false negative.

Second, consider the case of a false positive. Suppose that model A is the "true" model, i.e. the true cointegration rank is 1. Suppose we make a false positive in the second stage of the test procedure. That is, we infer that the cointegration rank is 2, add equation (8), and end up with model B, which is "biased". In this case the causal misspecification bias is $CE_B - CE_A$, which is equal and opposite to the previous case.

Furthermore, in addition to bias, there is the problem of infeasibility of causal effects. To see this, consider that the "true" model is model C, which has three cointegration equations. Because there are only three endogenous variables, the three cointegration equations give a unique solution for the variables. Because the values are determined by the equations, we can not intervene with them without creating a contradiction. Hence, the causal effects between the variables can not be calculated, i.e. they are infeasible. If we make a false negative in this case and leave out cointegration equations, the analysis of causal effects is in logical contradiction with the "true" model.

To conclude, because of the uneven treatment of false positive and negative, the common practice fails to deliver unbiased estimates of causal effects. Moreover, a failure to detect a full cointegration rank leads to a logically inconsistent causal analysis.

In section 4 we show how bias emerges with the CKC literature, because of a missing equation.

3. Carbon Kutznets curve

3.1. The new approach

We apply the causal framework to analyze a recent strand of literature in environmental economics dealing with the long-standing debate on the environmental Kuznets curve (EKC), which has received novel attention as concerns for climate change and the need for global mitigation action have gained more awareness. The EKC depicts a relationship between emissions and output: at low levels of economic development growth increases emissions, but at higher levels of output the relationship is reversed. Graphically this implies an "inverted U" shape for the function of output to emissions (see Figure 1). When the focus is particularly on carbon dioxide emissions, the relationship is referred to as the carbon Kuznets curve (CKC).

The CKC-hypothesis has substantial relevance to development and climate change mitigation policies. Under the CKC-hypothesis, economic growth would ultimately contribute to the reduction of emissions, implying synergy between development and mitigation policy goals. Typically the alternative is to assume that emissions grow as output grows. This would imply a conflict

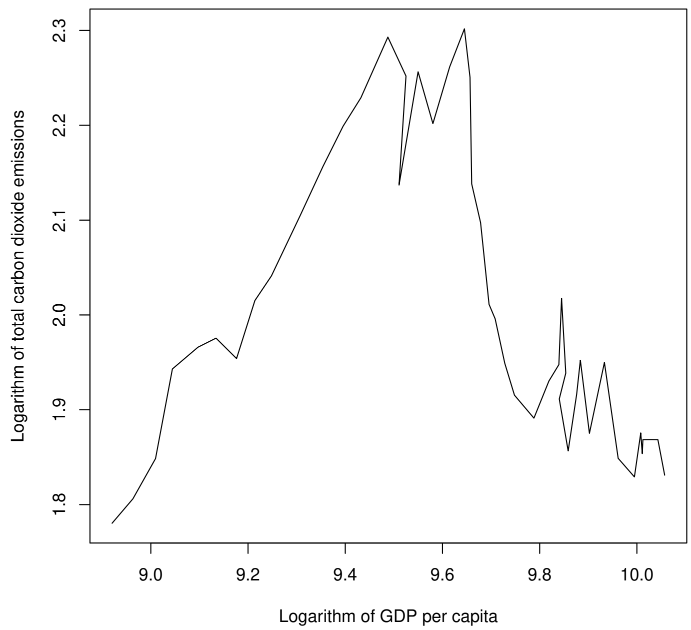


Figure 1: Relationship between the logarithm of per capita CO_2 emissions per capita and the logarithm of per capita GDP in France, 1960-2006.

between development and mitigation goals, and a need for separate climate policies.

Over the last few decades the EKC-hypothesis has generated an enormous amount of literature. The recent strand of literature has attempted to merge the CKC literature (emissions-output-nexus) with a related topic concerning the relationship between energy consumption and output (energy-output-nexus)(See Figure 2).[16]

[16]In a precursory study, Richmond and Kaufmann (2006) attempt to estimate the tipping point of the CKC with various model specifications. Some of these model specifications use the consumption shares of different fuel types to explain carbon dioxide emissions levels. The seminal work by Ang (2007) examines the relationship between emissions, energy consumption, and output in France using cointegration methods and a vector error-correction model (VECM). Total energy consumption is included as an explanatory variable to tackle omitted variable bias. Apergis and Payne (2009, 2010) extend

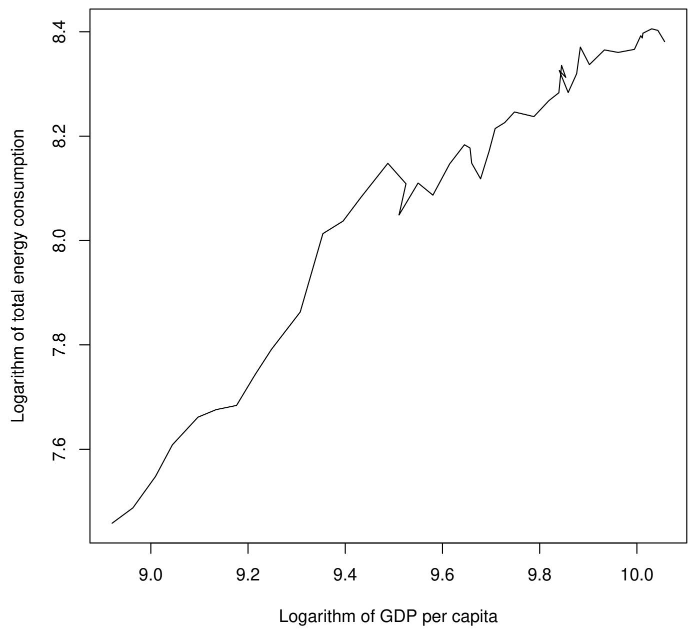


Figure 2: Relationship between the logarithm of total energy consumption per capita and the logarithm of per capita GDP in France, 1960-2006.

We describe three problems concerning the foundations of this new strand of literature. First, the nonlinearity of the CKC model is not compatible with vector autoregression (VAR) models, because it creates a binding yet neglected constraint for the model, which compromises the integrity of the

---

and apply this method for panel data on South American countries and for the countries of the Commonwealth of Independent States. Pao and Tsai (2010) applies this to panel data on BRIC countries (Brazil, Russia, India, and China) and later (2011) add foreign direct investment as a regressor. Soytas et al. (2007) use emissions, energy consumption, and output among others variables in a vector autoregression model (VAR) for the United States. Soytas and Sari (2009) apply a similar method for Turkey. Halicioglu (2009) adds foreign trade and uses an autoregressive distributed lag model (ARDL) model for Turkey. Jalil and Mahmud (2009) use an ARDL model for data on China and add foreign trade as an additional explanatory variable, while Jalil and Feridun (2011) add financial development to the equation.

estimators. The second and third problems relate to the inclusion of energy consumption as a explanatory variable.[17] The second problem arises because emissions are not measured directly in the datasets that are used by the referred articles. Carbon dioxide emissions are defined by a linear function of different fuel commodities. As a result, controlling for the level of energy use means that only carbon intensity of energy is allowed to vary. Subsequently the interpretations of the parameters are very different. The third problem is caused by the dependence between energy use and output. When this dependence is recognized, we see that the model is biased.

In this paper we discuss the problems related to the new strand of literature by focusing on the seminal work by Ang (2007). Some of the articles of the strand use slightly different methods and models so the problems manifest in different ways.[18] But they all have the common feature of controlling for energy in a CKC model. As we restrict our analysis to the article by Ang (2007), we can only cast serious doubt on the validity of the other papers and motivate a need for a re-evaluation. It is then the task of another paper to assess their viability.

3.2. Ang's model

Next we present the model introduced by Ang (2007). The CKC-hypothesis is examined using cointegration and vector error-correction modelling techniques (see e.g. Engle and Granger, 1987; Engle et al., 1989) applied to data on France between 1960 and 2000. The time series on emissions, energy use, and output are assumed to include stochastic trends, therefore many traditional time series methods are not applicable.

A long-run relationship between the time series can exist if stochastic trends are common to variables. A common stochastic trend implies that that there is a linear combination of the time series such that the combination is stationary. In which case, the time series are said to be cointegrated.

[17]The articles in question only briefly comment the rationale for doing this. Some argue that it helps to tackle omitted variable bias, but, how it would solve the endogeneity problem, is left without any justification or discussion. Nevertheless, this is not a trivial matter, and is the source of the second and third problem.

[18]For example, Apergis and Payne (2009, 2010) use a very similar methodology to Ang's. But Richmond and Kaufmann (2006) could face very different complications as they explain emissions with fuel proportions, not total energy use. Soytas et al. (2007) and Soytas and Sari (2009) use a time series technique known as the Toda-Yamamoto procedure, which does not explicate a long-run model, as do vector error-correction models.

Such a relationship is specified by Ang (2007) as a long-run (or steady-state) model

$$c_t = \beta_0 + \beta_1 e_t + \beta_2 y_t + \beta_3 y_t^2 + u_t, \tag{9}$$

where c_t is carbon dioxide emissions, e_t is total energy use, y_t is real GDP measured in local currency, all measured in per capita terms and converted into natural logarithms, and u_t is a stationary error term.[19]

As in a typical CKC-model, the square of output is included to capture the nonlinearity in the CKC. The CKC-hypothesis implies that parameter β_2 is positive and β_3 is negative to form an upside-down parabola. The novel feature is the included regressor e_t.

In addition to the long-run model, Ang (2007) studies the dynamic causal relationship between the time series by specifying a VAR model and the corresponding error-correction representation that incorporates equation (9). The VAR model describes how the variables vary, in the short-run, around the long-run model.

If Ang's (2007) long-run model is to be used for policy analysis, the model must be interpreted causally. That is, to answer the central question of the CKC-literature (how does output affect emissions), we must be able to answer counterfactuals statements (if output was increased, how would emissions behave). In other words, we need a causal model.[20]

To consider the policy question of the CKC-hypothesis, the long-run model can be represented as a causal model with additive interventions. Its structure consists of equation (9) and of equations determining the value of the exogenous variables e_t, y_t, and u_t. Carbon emissions c_t is chosen as endogenous in order to answer policy relevant questions regarding the CKC-hypothesis.[21] In this simple case, it seems that the carbon emissions c_t is the

[19]The long-run and short-run models are actually components of the same VECM. The error term of the long-run model corresponds to the error-correction term of the VECM. (Engle et al., 1989.)

[20]Note that, even though Ang obviously refers with "causality" to Granger causality, the model must be appended with traditional causality if it is to have any policy relevance.

[21]Note that the estimation process of the VECM does not imply any particular causal ordering between the variables. This choice is made only for argument's sake: emissions must be determined inside the model if we are to assess effects on them.

only endogenous variable. This can be depicted by a graph:

$$\begin{array}{ccc} e_t & \searrow & \\ y_t & \longrightarrow & c_t. \\ u_t & \nearrow & \end{array} \tag{10}$$

In Section 4 we show that this causal ordering is not the full story.

To sum up, Ang's model is policy relevant and attempts to address the CKC-hypothesis only if it is understood as a causal model where emissions are the effect of output. We claim that, if the model is interpreted in this manner, it can not answer the CKC-hypothesis. Note that we make no claims about causality in the real world. We are only interested about the consistency of causal statements with respect Ang's model.

The following section will introduce the tools needed to present the problems in Section 4.

3.3. The data and definitions

To consider the CKC literature, it is important to take into account how the carbon dioxide emissions data is produced in the datasets that are used in the literature.[22] Essentially, there are no actual measurements of carbon dioxide emissions. They are simply calculated from energy statistics (see Appendix C).

We define an important concept: *Carbon intensity* A_t is the average emissions rate of energy consumption.[23] Carbon intensity measures how much

[22]Ang (2007), Apergis and Payne (2009, 2010), Soytas et al. (2007), Soytas and Sari (2009), and Jalil and Mahmud (2009) use data from the World Bank's World Development Indicators (WDI) dataset, which in turn uses carbon dioxide emission data calculated by the U.S. Department of Energy's Carbon Dioxide Information Analysis Center (CDIAC) (Boden et al., 2009). In the CDIAC dataset carbon dioxide emissions are calculated from consumed quantities of different fuel commodities and cement manufacturing. The CDIAC dataset uses energy statistics by the United Nations Statistics Division (UNSD) among others. UNSD data is used for the time period analyzed in this paper. The WDI dataset uses energy statistics compiled by the International Energy Agency (IEA). Richmond and Kaufmann (2006) use data compiled by the IEA on energy use, and calculates the carbon dioxide emissions by multiplying fuel use by the appropriate carbon content factor.

[23]Note that here carbon intensity refers to the ratio of carbon emissions to energy consumption. This is not to be confused with carbon intensity of output which is the ratio of carbon emissions to output.

carbon dioxide emissions one unit of energy produces on average. Appendix C gives a formal definition.

Given the calculation form for carbon in the dataset and the concept of carbon intensity, we can derive identity

$$C_t \equiv E_t A_t + X_t, \tag{11}$$

where C_t is carbon dioxide emissions, E_t is total energy consumption, and X_t is emissions from gas flaring and cement manufacturing, all measured per capita. Appendix C shows how this identity is derived.

It is important to note, that this is simply an accounting identity derived from the definitions of the dataset, so it must be satisfied in the sample. That is, identity (11) holds by definition in the dataset.

To derive an algebraically more convenient form, we note that gas flaring and cement manufacturing amount only to a percent of total carbon emissions in the data, thus they can be omitted, i.e. set to zero. Therefore taking a natural logarithm of equation (11) gives

$$c_t = e_t + a_t, \tag{12}$$

where the variables are the corresponding logarithms of the capital letter variables.

Next we present the three problems in the new CKC literature and apply the causal framework.

4. The misspecifications in the new CKC literature

4.1. Transformations in a VAR model

The first problem arises because of the simple functional relationship between the observed variables y_t and y_t^2. Inspired by Haavelmo (1943), we notice that the system of equations implies a restriction between the joint distributions, which should have been taken into account in the estimation of the parameters. More specifically, the assumption of normally distributed i.i.d error terms (Johansen, 1988) is in contradiction with the VAR model equations.

To show this, lets look at the VAR model presented by Ang (2007):

$$x_t = a_0 + \sum_{i=1}^{p} A_i x_{t-i} + \varepsilon_t, \tag{13}$$

where $x_t = (c_t, e_t, y_t, y_t^2)$ is a vector of logarithms of observed variables, $\varepsilon_t \in \mathbb{R}^4$ is a vector of error terms for time $t = 0, \ldots, T$, $a_0 \in \mathbb{R}^4$ is a vector of constants, $A_i \in \mathbb{R}^{4\times 4}$ is a matrix of parameters for lag i, and p is the number of lags.

To simplify the analysis, we restrict to a model only two variables, y_t and y_t^2, and assume that $p = 1$, $a_0 = 0$, and $A_i = [a_{jk}] \in \mathbb{R}^{2\times 2}$. Hence our model consists of equations

$$y_t = a_{11} y_{t-1} + a_{12} y_{t-1}^2 + \varepsilon_t^1 \quad \text{and} \tag{14}$$

$$y_t^2 = a_{21} y_{t-1} + a_{22} y_{t-1}^2 + \varepsilon_t^2. \tag{15}$$

Now plugging y_t of equation (14) into equation (15) and rearranging gives

$$\varepsilon_t^2 = (a_{11} y_{t-1} + a_{12} y_{t-1}^2 + \varepsilon_t^1)^2 - a_{21} y_{t-1} - a_{22} y_{t-1}^2. \tag{16}$$

First,we notice that equation (16) constrains a polynomial relationship between error terms ε_t^1 and ε_t^2, hence they can not both be normally distributed, when the lagged variables are given at time t.

Second, we can show that the error terms are not independent over time. To see this, first note that equation (14) implies that $\frac{\partial y_{t-1}}{\partial \varepsilon_{t-1}^1} = 1$. Now using the chain rule and differentiating equation (16) gives

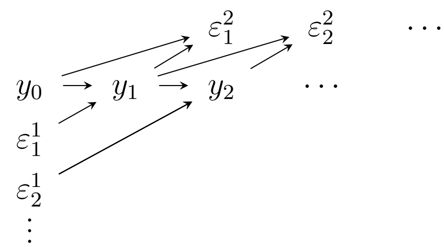
$$\frac{\partial \varepsilon_t^2}{\partial \varepsilon_{t-1}^1} = \frac{\partial \varepsilon_t^2}{\partial y_{t-1}} \frac{\partial y_{t-1}}{\partial \varepsilon_{t-1}^1} = 2(a_{11} y_{t-1} + a_{12} y_{t-1}^2 + \varepsilon_t^1)(a_{11} + 2a_{12} y_{t-1}) - a_{21} - 2a_{22} y_{t-1},$$

which is generally non-zero. Hence the value of error term ε_t^2 depends on ε_{t-1}^1, and therefore they can not be chosen independently.

This means that the assumption of normally distributed i.i.d error terms, which is required by Johansen's (1988) estimation method, can not be satisfied and the estimates are not reliable. We also see a much more general property: including a transformation of an endogenous variable into a VAR model as an endogenous variable creates an implicit constraint between error terms.

Alternatively, the problem can be seen easily by using the causal framework. First note that this VAR model specification does not constitute a structure if we assume that all error terms are exogenous. This is because, for each time period t, we have two model equations but only one endogenous variable, as y_t^2 is a simple transformation of y_t. That is, given that the error terms are exogenous, there are two model and two exogeneity equations for each t but only three variables, y_t, ε_t^1, and ε_t^2. This means that the regularity conditions are not satisfied and the equations do not form a structure, which implies that the model specification is inconsistent.

Instead, if we assume that one error term is endogenous, we get a consistent model. For this we can draw a causal graph that illustrates how the variables are determined:



Here we clearly see, that error term ε_1^1 is a cause of ε_1^2 and ε_2^2, which reveals the problematic dependency between the error terms. This is an example how the causal framework helps in detecting contradictions in the model setup.

Because there are other problems of misspecification in Ang's model, on which we want to focus, we assume in the subsequent sections that the aforementioned problem is only minor and does not change the estimates significantly.

4.2. The interpretation of the parameters

Now we turn to the second problem. Wrong interpretation of the parameters and resulting wrong conclusions arise from the definition of carbon emissions in the dataset (Section 3.3). It is worth emphasizing, that this problem is not about the fitness of the model, but is related to the specification of the model. We take Ang's model as given in order to deduce what it actually implies.

The long-run model equation (9) represents the policy relevant part of Ang's model. In relation to this equation causal and counterfactual statements are made (such as: if energy use would be higher, then emissions would be higher). Therefore it is fair to interpret the model equation (9) as a part of a causal model as described in Section 2. In Appendix D.2 we show that the problem described in this section also emerges when considering the short-run model, i.e. the VAR model.

In the context of CKC, the parameter of interest is the causal effect of output on carbon emissions. That is, we want to know how output effects emissions. In model equation (9), the proposed interpretation of the partial causal effect of y_t on c_t,

$$\frac{\partial c_t}{\partial y_t} = \beta_2 + 2\beta_3 y_t, \tag{17}$$

would be that it quantifies the causal relationship between emissions and output.[24] This would be the Marshallian ceteris paribus change that assumes other variables constant (Heckman and Vytlacil, 2007).

This however, does not take into account the conceptual dependence between energy and carbon emissions that is captured by identity (11). Recognizing this dependence reveals that the causal effect (17) has a much more narrow interpretation than implied.

Before applying the causal framework, we give, in more conventional means, three arguments for the existence of a problem, to show the necessity of a rigorous framework:

First, since the partial derivative (17) requires that total energy use e_t is held constant, we notice from identity (11) that, in this case, the level of carbon dioxide emissions C_t can only change through changes in carbon intensity A_t, or gas flaring and cement manufacturing emissions X_t. As the significance of X_t is negligible and assumed zero, like in equation (12), a change in output results in a change in carbon intensity a_t. The causal effect (17) can be interpreted only as the causal effect of output y_t on emissions c_t through carbon intensity a_t. This ignores the effect of y_t on c_t through energy use e_t. As a result, the model is actually a regression analysis of carbon intensity, instead of carbon emissions.

Second, the problem can be also seen by comparing the causal effect (17) and the derivative of identity (11). To simplify, assume gas flaring and cement manufacturing emissions are constants. Now partially derivating identity (11) with respect to y_t gives

$$\frac{\partial c_t}{\partial y_t} = e_t \frac{\partial a_t}{\partial y_t} + \frac{\partial e_t}{\partial y_t} a_t. \tag{18}$$

If emissions level e_t is held constant, as is required to calculate the causal effect (17), the second term on the right hand side of (18) is omitted. This

[24] To be exact, we are interested in the expected conditional partial derivative, but to ease notation, we do most of the analysis as if it was a deterministic model.

means that the causal effect (17), which Ang (2007) investigates, is only the first term in (18).

Third, a more explicit regression equation can be formulated. When the negligible effects of gas flaring and cement manufacturing are assumed to be zero, equation (12) can be plugged into equation (9) to eliminate c_t. Rearranging gives equation

$$a_t = \beta_0 + (\beta_1 - 1)e_t + \beta_2 y_t + \beta_3 y_t^2 + u_t. \tag{19}$$

Here we see that model equation (9) is equivalently an regression on carbon intensity a_t, and the functional form between a_t and output y_t is exactly the same as between carbon emissions c_t and y_t.

It easy to oppose this kind of reasoning with deceptively convincing counterarguments, by presenting different counterfactuals and relationships. To make a rigorous account of our argument and possible counterarguments, we use Simon's theory.

First, we assume gas flaring and cement manufacturing emissions are zero for convenience. Second, we form a set of model equations

$$c_t = \beta_0 + \beta_1 e_t + \beta_2 y_t + \beta_3 y_t^2 + u_t, \tag{20}$$
$$c_t = e_t + a_t, \tag{21}$$

where the first equation is (9), which is assumed to hold for arguments sake, and the second equation is (12), which result from identity (11) and the previous assumption. Third, we select the equations determining the value of the exogenous variables. This can be understood as fixing the inputs (see Heckman and Vytlacil, 2007). As exogenous we select variables e_t because it is defined as a control variable, y_t because it is represents our action, and u_t because it represents the other factors whose effect we want to omit as a ceteris paribus assumption. We represent these by equations

$$e_t = e_t^0, \quad y_t = y_t^0, \quad u_t = u_t^0, \tag{22}$$

where e_t^0, y_t^0, and u_t^0 are given values. We can think that the values are picked by some external mechanism, e.g. nature or the government.

Together equations (20)–(22) form a structure for which we can draw a graph. Solving the set of equations reveals the causal relationships between the variables, which can be depicted as

$$\begin{tikzcd} e_t \arrow[dr] \arrow[drr] & & \\ y_t \arrow[r] & c_t \arrow[r] & a_t. \\ u_t \arrow[ur] & & \end{tikzcd} \tag{23}$$

Regarding the whole structure, carbon intensity a_t is the ultimate endogenous variable. In other words, if a change in output y_t causes a change in emissions c_t, this means that carbon intensity a_t must have changed. This is not sound with the theory of the carbon Kutznets curve because lower carbon intensity should be a cause of lower emissions, not the other way around.

Next show how the framework can be used to tackle some intuitive counterarguments. What if we claim that e_t is not exogenous and leave equation $e_t = e_t^0$ out? This would not be a structure any more and variables e_t, c_t, and a_t would not be determined. What if then choose c_t to be exogenous instead of e_t? Then we would get a different causal ordering and a new graph, but this would not be relevant for inspecting the CKC hypothesis as we are interested in the effect of y_t on c_t.[25]

What if we claim that there is an additional relationship between the variables? This means adding a new equation to the structure. This can give us a different causal ordering and a new graph. First, if the new equation contains a new variable, the effect is similar to adding equation (12) to model equation (9), as done above. Graph (10) corresponds to model equation (9) alone, but adding equation (12) gives us graph (23). Second, if the new equation does not contain a new variable, we have to make one variable endogenous. This is what we propose in the next section.

As we argued that Ang's model is explaining carbon intensity instead of carbon emissions, one might wonder why the model fits so well. We address this question in Appendix E and argue the fit is misleading due to the polynomial shape of the model.

4.3. Bias

The third problem is a misspecification bias rising because energy use e_t is dependent on output y_t. To begin, let's look at the data. First, note that

[25] Also this is technically restricted because the binomial in (9) has two roots.

Figure 2 depicts levels of energy use over different output levels. It seems like there is a relationship between the variables: energy use is larger when output is larger. This suggest that producing more output requires more energy, as is apparent from theory. Second, note that, as identity (11) implies, variations in both carbon intensity and energy use are essential for the CKC. This can be seen from Figure 3. Here the development of carbon emissions in France (curve A) has been decomposed in to a growing energy consumption (B) and a declining carbon intensity(C).[26] This shows that, without the growth of energy consumption, emissions in 2006 would be 60% less compared to 1960. On the other hand, without the shift to cleaner fuels, emissions would be 150% higher in 2006. This means that clearly both factors need to be accounted for.

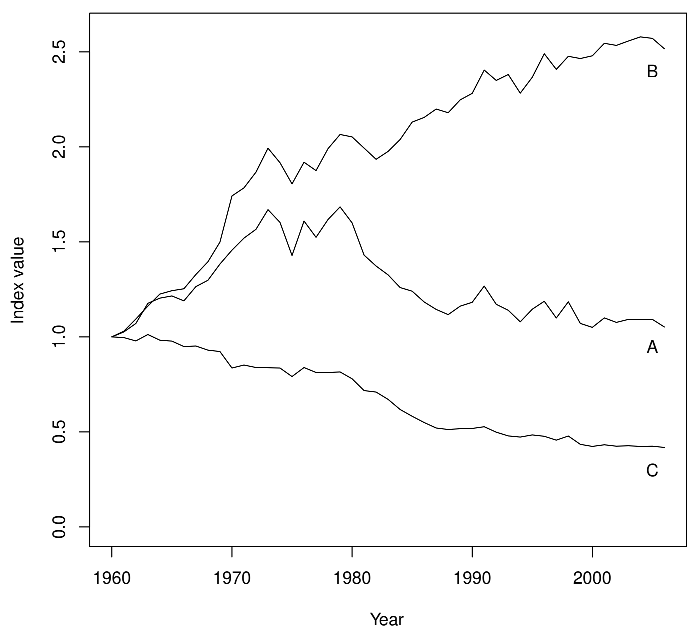


Figure 3: Curve A is the index of carbon emissions, B is a index energy consumption, and C is the carbon intensity. By definition $A = BC$.

[26]To be more specific, $A = \frac{c_t}{c_{1960}}$, $B = \frac{e_t}{e_{1960}}$, and $C = \frac{c_t}{c_{1960}} / \frac{e_t}{e_{1960}}$. Hence $A = BC$.

The misspecification problem can be rigorously described by using the causal framework presented in chapter 2. The long-run model is the base of policy implications, and counterfactuals are stated with the long-run model equation (9). Therefore we can interpret the long-run model as a causal model and use the framework to show inconsistencies in the way causal statements are made.

We begin by first noting the three mechanisms dictated by the CKC-hypothesis and our knowledge of the definitions. These mechanisms describe how the variables relate to each other and allow us to construct a structure that can be compared to Ang's structure.

First, we take into count the definition of carbon emissions in identity (11). To clarify the results, we omit emissions from gas flaring and cement manufacturing, and use equation (12). This states that carbon emissions can be decomposed into carbon intensity a_t and energy use e_t.

Second, output y_t is a cause of carbon intensity a_t according to the CKC-hypothesis. This is also implied by Ang (2007) as shown in the previous section. Because we want to assess the bias in relation to Ang's model, we assume that model equation (9) is satisfied.[27] But because it actually describes a mechanism for carbon intensity a_t (as shown in the previous section), we use the equivalent equation (19). In other words, we assume that equation (19) is satisfied to make a sensible comparison.[28]

Third, we note that also energy use e_t depends on output y_t. This basic notion, which is fairly evident (the details are the subject of the immense energy-output-nexus literature), is actually the motivation behind the strand of literature initiated by Ang (2007), and is essential to the CKC-hypothesis. Nonetheless it is unintentionally neglected due to the model formulation. To capture this relationship, we simply assume that there is a differentiable and monotonically increasing function e for which $e_t = e(y_t) + v_t$, where v_t is a error term.[29]

[27] Note that we do not argue that this model is empirically valid in all respects. To the contrary, as we argue in Appendix E, carbon intensity is not a parabola. We construct our model to isolate one only one faulty aspect of Ang's model, i.e. the bias, and focus on that.

[28] Note that also equation (9) could be chosen but this would result in an unfounded causal ordering (presented in the previous section) without affecting the bias.

[29] The existence of another cointegration equation can not be ruled out by the cointegration tests. Tests, that are derived from Johansen (1988) and used in Ang (2007), reject the null hypothesis of zero cointegration equations against the alternative of one *or more*

In addition, we fix y_t because it represents our action variable, and we fix u_t and v_t as ceteris paribus assumptions. This is denoted by equations $y_t = y_t^0$, $u_t = u_t^0$, and $v_t = v_t^0$.

These three mechanisms and three exogeneity equations form a set of equations, a structure with additive interventions,

$$c_t = a_t + e_t \tag{24a}$$
$$a_t = \beta_0 + (\beta_1 - 1)e_t + \beta_2 y_t + \beta_3 y_t^2 + u_t \tag{24b}$$
$$e_t = e(y_t) + v_t \tag{24c}$$
$$y_t = y_t^0 \tag{24d}$$
$$u_t = u_t^0 \tag{24e}$$
$$v_t = v_t^0. \tag{24f}$$

The causal ordering of the structure (24) can be depicted by graph

$$\begin{tikzcd}
v_t \arrow[r] & e_t \arrow[drr] \arrow[ddr] & & \\
y_t \arrow[ur] \arrow[drr] & & & c_t. \\
u_t \arrow[rr] & & a_t \arrow[ur] &
\end{tikzcd}$$

The reasoning in the graph can also be expressed less formally. First, suppose y_t, u_t, and v_t are determined by an external process, the economy for example. Now also e_t is determined by y_t and v_t through the mechanism (24c). When e_t, y_t, and u_t are known, using equation (24b), also a_t is determined. Now e_t and a_t are set, so carbon emissions c_t is known by definition with equation (24a).

Note that we could, for example, form the structure with model equations (9), (24a) and (24c)–(24f), but this would alter the causal ordering. As Simon (1953) show, this would be empirically indistinguishable from structure (24), even if it is theoretically invalid. In other words, the identification of the model is only partial. We choose the structure (24), from a set of empirically equivalent structures, because the resulting causal ordering is more reasonable. This, however, does not affect our main concern, the bias.

cointegration equations. (See e.g. Lütkepohl, 2005, p. 329; Watson, 1994; Hamilton, 1994)

To assess the bias, we need to calculate a causal effect for the model defined above, and compare it with the causal effect in Ang's model. The magnitude of the causal effect of output y_t on carbon emissions c_t in structure (24) can be calculated by applying the implicit function rule, as shown in Appendix F. The (total) causal effect is given by

$$\frac{\mathrm{d}c_t}{\mathrm{d}y_t} = (\beta_2 + 2\beta_3 y_t) + e'\beta_1. \tag{25}$$

Now causal effect (25) can be compared to the biased interpretation in expression (17). We see clearly, that the model specification of Ang (2007) is biased by the term $-e'\beta_1$, which is negative in the plausible case. First, e' is positive when larger output implies more energy use. Second, the parameter β_1 should also be positive, as energy use has positive effect on carbon emissions.

The negative bias has two implications for the shape of the CKC.

First, the tipping point of CKC is at a higher level of output when bias exists. We show this in Appendix F. This means the tipping point will occur later than estimated in Ang (2007). This is supported by Figure 1, where we can see that Ang's estimate for the tipping point, 9.31, is clearly on the left side of the parabola.

A second implication for the shape is that the unbiased CKC grows quicker and declines more slowly, than the biased one. This is simply due to the fact that, for all levels of output y_t the biased causal effect is smaller than the unbiased one. Before the tipping point of the biased CKC, carbon emissions are actually growing faster. After the biased CKC has tipped, emissions are actually still growing for awhile. And after the true tipping point, emissions are declining but slower then the biased CKC implies.

The unbiased shape draws a more pessimistic view regarding the conflict between development and climate change policy goals.

5. Conclusions

We have defined a framework for analyzing causality in econometric models, and proposed a solution to key controversies in the literature of causal analysis. The framework allows to rigorously analyze causal implications in complex models where analysis has previously not been tractable.

We have shown that a common practice of policy analysis with VECM can result in biased or even infeasible estimates due to an unsuited testing procedure. If VECMs are to be used in policy evaluation, the testing procedure for the cointegration rank should avoid false positives and negatives equally, and absolve possible logical contradictions. On the other hand, it might be more fruitful to approach the issue from estimation theory instead of testing.

Regarding the CKC literature, we have shown, first, that using a transformation of an endogenous variable as an endogenous variable in a VAR model creates a constraint. Second, neglecting the dataset definitions alters the interpretation of the model parameters significantly. As a result, the new CKC literature answers a very different question compared to conventional CKC-literature. The estimated relationship is not the CKC as a whole. For the most part, it just estimates the relationship between carbon intensity and output, which neglects the causal effect through energy use. Third, energy use causes a misspecification bias, when a change in output implies a change in the level of energy use. As a result, the criticized model specification gives an overly optimistic view of the compatibility of development and environmental policy goals. If there is a tipping point, it occurs later than expected. Before tipping, output increases emissions faster, and afterwards, emissions drop slower than anticipated. To answer any relevant questions about the CKC-hypothesis, one can not simply combine the energy-output and carbon-output nexuses into one equation.

Appendixes

Appendix A. Proofs

Lemma 1. Substructures of a structure are disjoint.

Proof. Assume that the proposition is not true. Now let A and B, $A \neq B$, be proper substructures of a structure, such that they have common equations, i.e $E = A \cap B$ is non-empty. Let $S = A \cup B$. Denote by N_A, N_B, N_E, and N_S the number of equations, and by n_a, n_b, n_e, and n_s, the number of non-trivial variables in sets A, B, E, and S, respectively.

As a rule of set theory, equations

$$N_S = N_A + N_B - N_E \quad \text{and} \tag{A.1}$$

$$n_s = n_a + n_b - n_e \tag{A.2}$$

must be satisfied.

Substructures must contain as many equations as variables, i.e. $N_A = n_a$ and $N_B = n_b$. The second regularity condition requires that $n_s \geq N_S$. Now plugging these into (A.1) gives inequality $n_s \geq n_a + n_b - N_E$. Plugging this into equation (A.2) and rearranging gives $n_e \leq N_E$. By the second regularity condition, we must have $n_e \geq N_E$. Now $n_e = N_E$. According to the third regularity condition E must be solvable. Due the definition of a substructure, B and C should not contain solvable subsets, so subset E should not be solvable, i.e. we have a contradiction. □

Lemma 2. The partition is unique.

Proof. Assume that there is an alternative partition. Now there must be a substructure in the alternative partition that, first, contains a equation that is in a substructure of original partition, and, second, contains another equation that is not in substructure of original partition. Therefore these substructures are not disjoint, which contradicts Lemma 1. □

Appendix B. Calculating the causal effect

Consider a causal model with continuous variables and differentiable model functions g with additive interventions that define deep model function $f\colon \mathbb{R}^n \to \mathbb{R}^n$. Let

$$h(y, x, a, x^0, a^0) \equiv \begin{bmatrix} f(y, x, a) \\ x - x^0 \\ a - a^0 \end{bmatrix} = \begin{bmatrix} g(y, x + a) \\ x - x^0 \\ a \end{bmatrix} = 0$$

be a structure of the corresponding causal model.

Let J_h denote the Jacobian determinant of h derivated over $v = (y, x, a)$, that is

$$J_h = \begin{vmatrix} \frac{\partial h_1}{\partial v_1} & \frac{\partial h_1}{\partial v_2} & \cdots & \frac{\partial h_1}{\partial v_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial v_1} & \frac{\partial h_n}{\partial v_2} & \cdots & \frac{\partial h_n}{\partial v_n} \end{vmatrix}$$

and let $J_h(i, j)$ denote a corresponding determinant where ith row has been replaced by the derivate $\frac{\partial h}{\partial a_j^0}$, that is

$$J_h(i, j) = \begin{vmatrix} \frac{\partial h_1}{\partial v_1} & \cdots & \frac{\partial h_1}{\partial v_{i-1}} & \frac{\partial h_1}{\partial a_j^0} & \frac{\partial h_1}{\partial v_{i+1}} & \cdots & \frac{\partial h_1}{\partial v_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial v_1} & \cdots & \frac{\partial h_n}{\partial v_{i-1}} & \frac{\partial h_n}{\partial a_j^0} & \frac{\partial h_n}{\partial v_{i+1}} & \cdots & \frac{\partial h_n}{\partial v_n} \end{vmatrix}.$$

Now the causal effect is defined by equation $\frac{\mathrm{d}v_i}{\mathrm{d}a_j^0} = -\frac{J_h(i,j)}{J_h}$.

With some elementary but heavy algebra, one can show that $\frac{\mathrm{d}v_i}{\mathrm{d}a_j^0} = \frac{\mathrm{d}v_i}{\mathrm{d}x_j}$, where the latter derivate is calculated from set of equations $g(y, x) = 0$. This means, that in the additive interventions case causal effects can be calculated simply from the model equations.

Appendix C. Definition of carbon emissions

In this appendix we derive a identity from the definitions of the dataset. This identity is the source the second problem presented in Section 4.

First, the dataset in use (WDI) defines carbon dioxide emissions as a linear function of fossil fuel combustion and cement manufacturing. The amount of carbon dioxide emissions caused by combustion is determined by the chemical composition of the fuel. The emitted amount of carbon dioxide is calculated by multiplying the amount of fuel usage by a constant factor prescribed by the chemical properties of the fuel. Thus, the total carbon dioxide emissions C_t is a linear combination of the usage of oil E_t^{oil}, solid fuels E_t^{solid}, natural gas E_t^{gas}, and gas flaring E_t^{flare}, in addition to emissions from cement manufacturing S_t, all measured in per capita term. More formally, that is,

$$C_t \equiv \alpha_{oil} E_t^{oil} + \alpha_{solid} E_t^{solid} + \alpha_{gas} E_t^{gas} + \alpha_{flare} E_t^{flare} + S_t, \tag{C.1}$$

where α_{oil}, α_{solid}, α_{gas}, $\alpha_{flare} > 0$ are the related ratios of emissions to fuel quantity. (See Boden et al., 2009)

Second, total energy use E_t can be defined as the sum of oil E_t^{oil}, solid fuels E_t^{solid}, natural gas E_t^{gas}, and other energy sources E_t^{other}, such as nuclear energy and renewable fuels, which do not cause emissions in the aforementioned sense. Gas flaring does not result in energy production. Therefore

$$E_t \equiv E_t^{oil} + E_t^{solid} + E_t^{gas} + E_t^{other}.$$

To clarify the notation we define two sets of variable: the set of energy commodities affecting carbon dioxide emissions, $\mathfrak{C} = \{oil, solid, gas, flare\}$, and the set of energy commodities that amount to total energy use, $\mathfrak{E} = \{oil, solid, gas, other\}$.

Next let's define the proportions of fuel commodities in terms of total energy use,

$$q_t^i \equiv \frac{E_t^i}{E_t},$$

where $q_t^i \geq 0$ for all $i \in \mathfrak{E}$ and $\sum_{i \in \mathfrak{E}} q_t^i = 1$ for any t. By rearranging and plugging this into identity (C.1) to eliminate E_t^i for each i, we get

$$C_t \equiv E_t \sum_{i \in \mathfrak{C} \cap \mathfrak{E}} q_t^i \alpha_i + \alpha_{flare} E_t^{flare} + S_t.$$

By interpreting the sum term as the average emissions rate of energy consumption, we can identify it as *carbon intensity* and denote it by A_t, so that

$$C_t \equiv E_t A_t + \alpha_{flare} E_t^{flare} + S_t. \tag{C.2}$$

This is simply an accounting identity derived from the definitions of the dataset, so it must be satisfied by the observed values.

To derive an algebraically more convenient form, we note that gas flaring and cement manufacturing amount only to a percent of total carbon emissions in the data, thus they can be omitted, i.e. set to zero. Therefore taking a natural logarithm of equation (C.2) gives

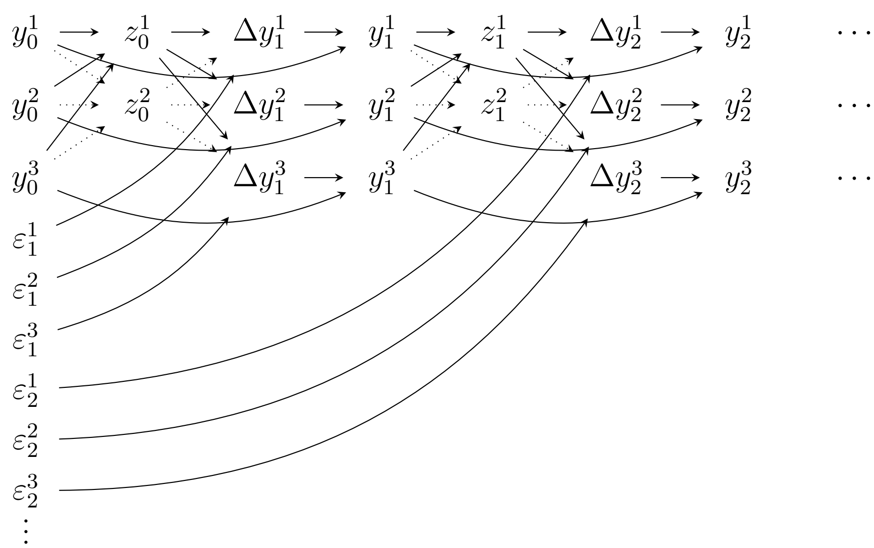
$$c_t = e_t + a_t,$$

where the variables are the corresponding logarithms of the capital letter variables.

Appendix D. Short run models

Appendix D.1. Causal ordering of VECMs

Model B in Section 2.6 consists of equations (5)–(8) and forms a structure which can be solved to get graph



(D.1)

when the values y_0^i are given.

Model A replaces equation (8) with $z_0^2 = 0$. The resulting graph equals (D.1) when the dotted arrows are removed and variables z_t^2 are put in the first row. Model A lacks the causal relationships represented by the dotted arrows in graph (D.1) and therefore has a different causal effects.

Appendix D.2. Problem 2 in the short-run model

Consider the VAR model equation (13) for time $t = 0, 1, \ldots, T$. To avoid the problem described in Section 4.1 we omit the equation describing y_t^2. That is,

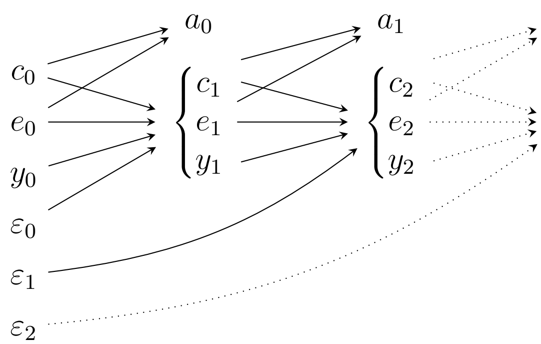
$$x_t = a_0 + \sum_{i=1}^{p} A_i x_{t-i} + \varepsilon_t, \quad t = 0, 1, \ldots, T,$$

where $x_t = (c_t, e_t, y_t)$ is a vector of logarithms of observed variables, $a_0 = (a_0^c, a_0^e, a_0^y)$ is a vector of constants, $A_i = [A_i^c \; A_i^e \; A_i^y]$ is a matrix of parameters for lag i, and $\varepsilon_t = (\varepsilon_t^c, \varepsilon_t^e, \varepsilon_t^y)$ is a vector of error terms.

We supplement this with equation (12) for each t.

Our exogeneity equations are $c_0 = c_0^0$, $e_0 = e_0^0$, $y_0 = y_0^0$, and $\varepsilon_t = \varepsilon_t^0$ for all $t = 0, 1, \ldots, T$.

Now we have $4(T+1))$ model equations and $T+4$ exogeneity equations and a equal amount of variables, $5T+5$, forming a structure. Now we get graph



,

which shows the same problematic causal ordering as in the long-run model.

Appendix E. Explaining the fit

In this appendix we consider Ang's model's relation to the data. In the section 4.2 we have assumed that Ang's model is correct (only the interpretation is wrong) and disregarded the problem shown in Section 4.1. So how can we account for the good fit of the model? As we have shown, in Ang's model changes in carbon emissions result from changes in carbon intensity. Therefore we look into this variable.

Carbon intensity a_t has decreased in France (Figure E.4) because of a decline in the share of heavily polluting fuels like coal. They have been replaced or outgrown by the use of oil, natural gas and nuclear energy (Kaufmann, 1992). Especially in the case of France, is seems that nuclear energy has had a significant impact (see Iwata et al., 2010).

This is in line with the traditional CKC-hypothesis: the decline in carbon intensity is a result of economic growth, however this is contrasted by a growth in energy use, resulting in an "inverted U"-shape between emissions and output.

Because model equation (9) is equivalent with (19), we should see an "inverted U"-shape also in carbon intensity. Figure E.4 does not suggest such a relationship for France in the period 1960-2006. But why then does Ang get significant estimates?

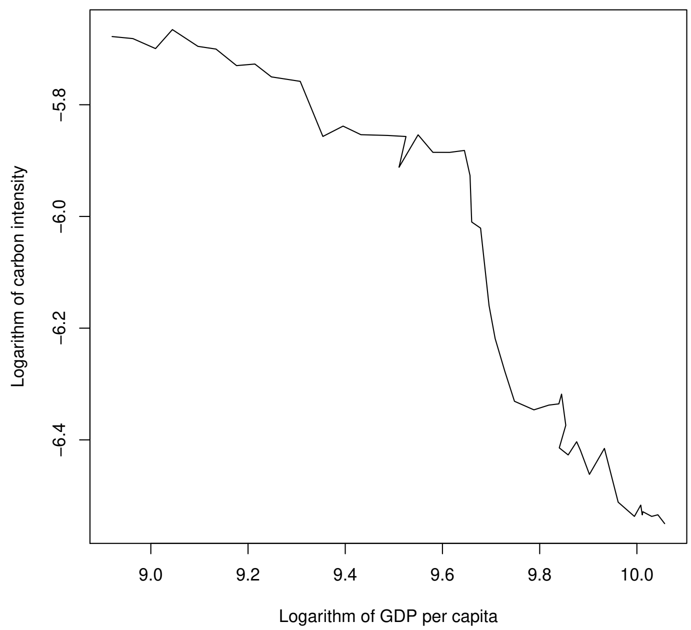


Figure E.4: The logarithm of Carbon intensity of energy use, measured in kilograms of carbon dioxide emissions per kilogram of oil equivalent energy, and the logarithm of per capita GDP.

The delusive fit of the parabola can be explained by inspecting the fitted model in Ang (2007). Ang reports the parameter estimates $\beta_0 = -161.38$, $\beta_1 = 2.25$, $\beta_2 = 31.11$, and $\beta_3 = -1.67$. By using these in model (9), rearranging, and taking the conditional expected value we get equation

$$c_t - 2.25e_t + 161.38 = 31.11y_t - 1.67y_t^2. \tag{E.1}$$

The left hand side of equation (E.1) is the part that is alloted for output to explain. Because β_3 is negative, the value of the right hand side as a function of y_t is an upside-down parabola. Suppose this is truly so. Then, to satisfy equation (E.1), also the left hand side should be a parabola.

The values of the left side terms are plotted against y_t in Figure E.5 along with the regression curve (dashed curve), i.e the right hand side of equation (E.1).[30] The misleading goodness of the fit of the “inverted U”-shaped curve found by Ang (2007) can be explained by observing, that most of the observations lie on the right hand side of the parabola.

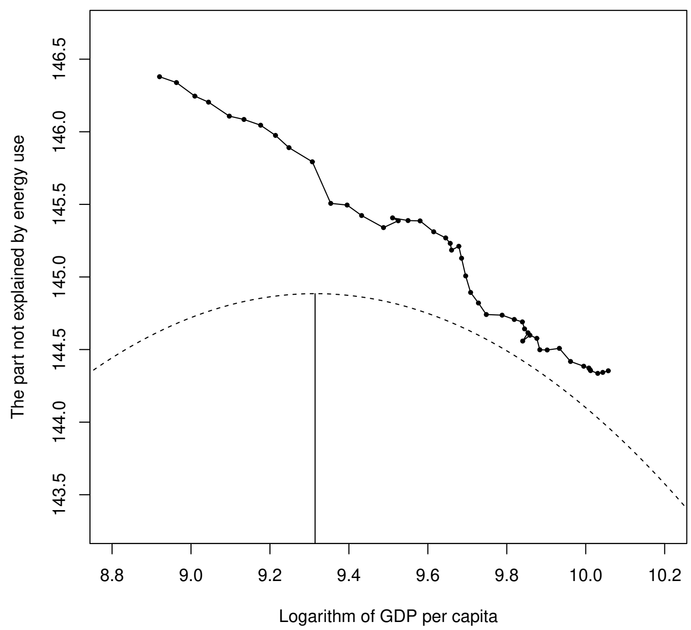


Figure E.5: The solid line is the relationship between the logarithm of per capita CO_2 emissions unexplained by energy use and the logarithm of per capita GDP. The dashed line is the corresponding regression curve, i.e. the right hand side of equation (E.1), and the vertical line indicates the tipping point.

However, this does not give a reason to suspect the existence of a tipping point in the carbon intensity or an "inverted U"-shape. As in Ang’s model,

[30] The seemingly biased fit in Figure E.5 might be due strong sensitivity to the rounding of the parameter values. Alternative parameter that have the same rounded value can shift the regression curve to give an opposite bias. On the other hand, this might be due to problem described in Section 4.1.

there exists a CKC if and only if the plot in Figure E.5 is polynomial, we can not arrive to our result simply by assuming it. That is, if we can not justify Figure E.5 (or Figure E.4) representing a polynomial shape (with a tipping point in the sample), we can not justify the existence of a carbon Kuznets curve.

Why the do we see a parabola shape in Figure 1? Here energy use is not fixed, thus the parabola is formed by the rising energy use and the declining carbon intensity. But in Ang's model the term $31.11 y_t - 1.67 y_t^2$ plots a curve for a fixed energy level. Therefore this term cannot be plotted in Figure 1, without letting energy use vary. This connection between output and energy use is the source of the problem in Section 4.3.

Appendix F. Mathematical derivations for section 5.3.

The magnitude of the causal effect of output y_t on carbon emissions c_t in structure (24) with additive interventions can be assessed by applying the implicit function theorem to get the total derivative

$$\frac{\mathrm{d}c_t}{\mathrm{d}y_t} = -\frac{\begin{vmatrix} 0 & -1 & -1 \\ -(\beta_2 + 2\beta_3 y_t) & 1 & 1-\beta_1 \\ -e' & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -1 & -1 \\ 0 & 1 & 1-\beta_1 \\ 0 & 0 & 1 \end{vmatrix}},$$

where we denote $e' = \frac{\partial e_t}{\partial y_t}$. By calculating the determinants, we get expression

$$-\frac{(\beta_2 + 2\beta_3 y_t) - e'(-1 + (1-\beta_1))}{-1},$$

which can be simplified to determine the (total) causal effect

$$\frac{\mathrm{d}c_t}{\mathrm{d}y_t} = (\beta_2 + 2\beta_3 y_t) + e'\beta_1. \tag{F.1}$$

Now causal effect (F.1) can be compared to the biased interpretation in expression (17). We see that the model specification of Ang (2007) is biased by the term $-e'\beta_1$, which is negative in the plausible case.

Next we show that the tipping point of CKC is at a higher level of output when bias exists. In the unbiased case the tipping point y_t^* is such that the causal effect (25) equals zero. This is equivalent to

$$y_t^* = \frac{-\beta_2 - e'\beta_1}{2\beta_3}.$$

Similarly, in the biased case the tipping point y_t^{**} satisfies

$$y_t^{**} = \frac{-\beta_2}{2\beta_3}.$$

Now, when $e'\beta_1 > 0$, adding β_2 to both sides gives $\beta_2 + e'\beta_1 > \beta_2$. Because β_2 is positive and β_3 is negative according to the CKC-hypothesis, we see that

$$\frac{-\beta_2 - e'\beta_1}{2\beta_3} > \frac{-\beta_2}{2\beta_3}.$$

By noting the tipping points, we get

$$y_t^* = \frac{-\beta_2 - e'\beta_1}{2\beta_3} > \frac{-\beta_2}{2\beta_3} = y_t^{**}.$$

That is, the true tipping point occurs at a higher level of output.

References

Aldrich, J., 1989. Autonomy. Oxford Economic Papers 41, pp. 15–34.

Ang, J.B., 2007. CO_2 emissions, energy consumption, and output in France. Energy Policy 35, 4772–4778.

Angrist, J.D., Krueger, A.B., 1999. Chapter 23 empirical strategies in labor economics, Elsevier. volume 3, Part 1 of *Handbook of Labor Economics*, pp. 1277–1366.

Apergis, N., Payne, J.E., 2009. CO_2 emissions, energy usage, and output in Central America. Energy Policy 37, 3282–3286.

Apergis, N., Payne, J.E., 2010. The emissions, energy consumption, and growth nexus: Evidence from the Commonwealth of Independent States. Energy Policy 38, 650–655.

Basmann, R.L., 1963a. The causal interpretation of non-triangular systems of economic relations. Econometrica 31, pp. 439–448.

Basmann, R.L., 1963b. On the causal interpretation of non-triangular systems of economic relations: A rejoinder. Econometrica 31, pp. 451–453.

Boden, T., Marland, G., Andres, R., 2009. Global, Regional, and National Fossil-Fuel CO_2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Dash, D., Druzdzel, M.J., 2008. A note on the correctness of the causal ordering algorithm. Artificial Intelligence 172, 1800–1808.

Demetrescu, M., Lütkepohl, H., Saikkonen, P., 2009. Testing for the cointegrating rank of a vector autoregressive process with uncertain deterministic trend term. Econometrics Journal 12, 414–435.

Druzdzel, M., Simon, H., 1993. Causality in bayesian belief networks, in: In Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence (UAI–93, Morgan Kaufmann Publishers, Inc. pp. 3–11.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: Representation, estimation, and testing. Econometrica 55, pp. 251–276.

Engle, R.F., Granger, C.W.J., Hallman, J.J., 1989. Merging short-and long-run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting. Journal of Econometrics 40, 45–62.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–38.

Haavelmo, T., 1943. The statistical implications of a system of simultaneous equations. Econometrica 11, pp. 1–12.

Haavelmo, T., 1944. The probability approach in econometrics. Econometrica 12, pp. iii–vi+1–115.

Halicioglu, F., 2009. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. Energy Policy 37, 1156–1164.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press. 1 edition.

Heckman, J.J., 2005. The scientific model of causality. Sociological Methodology 35, 1–97.

Heckman, J.J., 2008. Econometric causality. International Statistical Review 76, 1–27.

Heckman, J.J., 2010. Building bridges between structural and program evaluation approaches to evaluating policy. Journal of Economic Literature 48, 356–98.

Heckman, J.J., Vytlacil, E.J., 2007. Chapter 70 econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation, Elsevier. volume 6, Part 2 of *Handbook of Econometrics*, pp. 4779–4874.

Holland, P.W., 1986. Statistics and causal inference. Journal of the American Statistical Association 81, pp. 945–960.

Hoover, K.D., 2003. Review: Causality by Judea Pearl. The Economic Journal 113, pp. F411–F413.

Hoover, K.D., 2008. causality in economics and econometrics, in: Durlauf, S.N., Blume, L.E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke.

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47, 5–86.

Iwata, H., Okada, K., Samreth, S., 2010. Empirical study on the environmental Kuznets curve for CO_2 in France: The role of nuclear energy. Energy Policy 38, 4057–4063.

Jalil, A., Feridun, M., 2011. The impact of growth, energy and financial development on the environment in China: A cointegration analysis. Energy Economics 33, 284–291.

Jalil, A., Mahmud, S.F., 2009. Environment Kuznets curve for CO_2 emissions: A cointegration analysis for China. Energy Policy 37, 5167–5172.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12, 231–254.

Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press.

Kaufmann, R.K., 1992. A biophysical analysis of the energy/real GDP ratio: implications for substitution and technical change. Ecological Economics 6, 35–56.

Leamer, E.E., 1985. Vector autoregressions for causal inference? Carnegie-Rochester Conference Series on Public Policy 22, 255–304.

Leroy, S.F., 2001. A review of Judea Pearl's Causality. Journal of Economic Methodology 9, 100–103.

Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer.

Marschak, J., 1950. Statistical inference in economics, in: Statistical Inference in Dynamic Economic Models, John Wiley and Sons.

Matzkin, R.L., 2007. Chapter 73 nonparametric identification, Elsevier. volume 6, Part 2 of *Handbook of Econometrics*, pp. 5307–5368.

Mäki, U., 2008. scientific realism and ontology, in: Durlauf, S.N., Blume, L.E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke.

Mäki, U., 2011. Models and the locus of their truth. Synthese 180(1) , 47-63.

Neuberg, L.G., 2003. Causality: Models, reasoning, and inference, by Judea Pearl, Cambridge University Press, 2000. Econometric Theory 19, 675–685.

Neyman, J., 1935. Statistical problems in agricultural experimentation. Supplement to the Journal of the Royal Statistical Society 2, pp. 107–180.

Pao, H.T., Tsai, C.M., 2010. CO_2 emissions, energy consumption and economic growth in BRIC countries. Energy Policy 38, 7850–7860. Special Section: Carbon Reduction at Community Scale.

Pao, H.T., Tsai, C.M., 2011. Multivariate granger causality between CO_2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. Energy 36, 685–693.

Pearl, J., 1993. Bayesian analysis in expert systems: Comment: Graphical models, causality and intervention. Statistical Science 8, pp. 266–269.

Pearl, J., 2003. Comments on Neuberg's review of Causality. Econometric Theory 19, 686–689.

Pearl, J., 2009. Causality: Models, Reasoning and Inference. Cambridge University Press, New York, NY, USA.

Richmond, A.K., Kaufmann, R.K., 2006. Is there a turning point in the relationship between income and energy use and/or carbon emissions? Ecological Economics 56, 176–189.

Rubin, D.B., 1974. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. Journal of Educational Psychology .

Rubin, D.B., 1978. Bayesian inference for causal effects: The role of randomization. The Annals of Statistics 6, pp. 34–58.

Simon, H., 1953. Causal ordering and identifiability, in: Hood, W., Koopmans, J. (Eds.), Studies in econometric method. New York, NY: Wiley.. chapter 3.

Simon, H.A., Iwasaki, Y., 1988. Causal ordering, comparative statics, and near decomposability. Journal of Econometrics 39, 149–173.

Simon, H.A., Rescher, N., 1966. Cause and counterfactual. Philosophy of Science 33, 323–340.

Soytas, U., Sari, R., 2009. Energy consumption, economic growth, and carbon emissions: Challenges faced by an EU candidate member. Ecological Economics 68, 1667–1675. Eco-efficiency: From technical optimisation to reflective sustainability analysis.

Soytas, U., Sari, R., Ewing, B.T., 2007. Energy consumption, income, and carbon emissions in the United States. Ecological Economics 62, 482–489.

Strotz, R.H., Wold, H.O.A., 1960. Recursive vs. nonrecursive systems: An attempt at synthesis (part I of a triptych on causal chain systems). Econometrica 28, pp. 417–427.

Strotz, R.H., Wold, H.O.A., 1963. The causal interpretability of structural parameters: A reply. Econometrica 31, pp. 449–450.

Watson, M.W., 1994. Chapter 47 vector autoregressions and cointegration, Elsevier. volume 4 of *Handbook of Econometrics*, pp. 2843–2915.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. volume 1 of *MIT Press Books*. The MIT Press.