What is macroeconomic causality?[1]

Abstract (149 words): The paper aims to develop an adequate account of macroeconomic causality. It discusses the definition that is central to Woodward's interventionist account and the definitions that can be extracted from Hoover's remarks on privileged parameterization and from the potential outcome approach that Angrist and Kuersteiner have introduced into macroeconomics more recently. The definition to be defended can be regarded as the gist that is common to all three definitions when they are relieved of overly restrictive conditions. It says (roughly) that *X* causes *Y* iff there is a possible intervention on *X* that changes *Y*, where *X* and *Y* stand for macroeconomic aggregates, where an intervention is understood as a manipulation of an intervention variable *I* that satisfies conditions requiring that *I* be a cause of *X*, and that there be no confounders of *X* and *Y*, and where an intervention variable is either a variable or a parameter.

1. Introduction

Kevin Hoover's sophisticated work *Causality in Macroeconomics* begins by stating an undeniable truth: the ultimate justification for the study of macroeconomics is to provide secure knowledge on which to base policy; policy is about influencing outcomes, about control or attempted control; and the study of the particular connections that permit control of one thing to influence another is the study of causality (cf. Hoover, 2001, p. 1). If the study of macroeconomics is the study of the causality of the relations that macroeconomists believe obtain between aggregate quantities like inflation, GDP etc., then a question of significant philosophical and methodological interest relates to the nature of the causality of these relations. But a striking fact about that question is that macroeconomists or econometricians hardly pay any attention to it.

There are, as far as I can see, only four accounts of macroeconomic causality that have been proposed in response to that question. Clive Granger (1980, pp. 330, 336) is the only econometrician who comes up with an explicit definition of "cause". Nancy Cartwright (2009, p. 415) notes that David Hendry sometimes suggests that causality is super exogeneity. Perhaps the most elaborate account of macroeconomic causality can be found in Hoover (2001, 2011, 2013). And probably the most recent account can be reconstructed from the potential outcome approach that Joshua Angrist and Guido Kuersteiner (2011) apply to test for the causal effects of monetary policy shocks.

[1] I would like to thank Kevin Hoover, Winfried Pohlmeier, Julian Reiss, Wolfgang Spohn, my colleagues at the Center for Philosophy of Science of the University of Pittsburgh (which I visited in the fall 2016), and two anonymous referees for valuable comments and suggestions.

The aim of the present paper is not to speculate about the reasons why macroeconomists and econometricians tend to refrain from confronting the question of the nature of the causality of the relations that they think obtain between aggregate quantities.[2] The aim of the paper is to develop an adequate account of macroeconomic causality. In order to develop that account, I'm not going to discuss the Granger and super exogeneity accounts. Both accounts have received a lot of attention in the econometric literature, and Hoover criticizes them in particularly concise ways. He argues that Granger causality is neither necessary nor sufficient for macroeconomic causality (cf. Hoover, 2001, pp. 151-155), and that super exogeneity is not a necessary condition for macroeconomic causality because super exogeneity and causality come apart in the case of macroeconomic models with nonlinear cross-equation restrictions (cf. Hoover, 2001, pp. 167-8).[3] Both criticisms reverberate in what econometricians say about these accounts: G. S. Maddala and Kajal Lahiri (2009, p. 390) observe that "Granger causality is not causality as it is usually understood"; and in their seminal paper on exogeneity, Robert Engle, David Hendry, and Jean-François Richard (1983, p. 384) point out that "super exogeneity is a sufficient but not a necessary condition for valid inference under policy interventions".

In order to develop an adequate account of macroeconomic causality, I'm going to discuss Hoover's and Angrist and Kuersteiner's accounts. I will, in addition, analyze the extent to which James Woodward's interventionist account of causality can be applied to macroeconomics. Woodward (2003, pp. 18, 258, 321) acknowledges on numerous occasions that his account is firmly rooted in the pioneering work of some of the early econometricians. With the exception of Hoover (cf. e.g. 2011, pp. 339-343; 2013, pp. 45-55), however, macroeconomists have largely ignored that account (or work that is similar in spirit, such as Judea Pearl's). I'm going to argue against all three accounts that they are too restrictive: that Woodward's account makes reference to a condition that doesn't need to be satisfied in the case of macroeconomics, that Hoover unnecessarily restricts the notion of an intervention to interventions on parameters, and that Angrist and Kuersteiner's account relies on at least two conditions that don't need to be satisfied in the case of macroeconomics.

The account that I will defend can be regarded as the gist that is common to all three accounts when they are relieved of these overly restrictive conditions. The definition that is central to that account says that X directly type-level causes Y if and only if there is a possible intervention on X that changes Y while all causal parents of Y excluding X (i.e. all variables that directly type-level cause Y except X) remain fixed by intervention, where X and Y stand for macroeconomic

[2] Hoover (2004, p. 161) mentions one such reason when stating that this tendency "is a corollary to the rise of mathematical formalism – especially the dominance of Walrasian general-equilibrium models".

[3] Hoover (2001, pp. 167-8) refers to the same class of models to show that the faithfulness condition that Spirtes, Glymour, and Scheines (1993, p. 56) say needs to be satisfied if the structure of a causal graph is supposed to be inferred from the screening relationships discoverable in the data is likely to be violated in macroeconomics. For an example of a macroeconomic model with a nonlinear cross-equation restriction consider the rational expectations model to be presented in section 2 below.

aggregates, where an intervention is understood as a manipulation of an intervention variable I that satisfies conditions requiring that I be a cause of X, that manipulating I break all arrows directed into any parent of Y except X (i.e. into any variable other than X that directly type-level causes Y), and that there be no confounders (i.e. causes of both X and Y), and where an intervention variable is either a causal structure variable or a parameter.

I will begin by introducing the notion of a causal model and by presenting macroeconomic models that can be regarded as instances of that model (section 2). I will then discuss a macroeconomic application of the definition that is central to Woodward's interventionist account of causality (section 3), Hoover's definition of "causality" in terms of privileged parameterization (section 4), and a definition that can be extracted from the potential outcome approach that Angrist and Kuersteiner (2011) have introduced into macroeconomics more recently (section 5). In a final section 6, I will summarize the arguments of sections 3-5 and spell out in detail what I think is an adequate definition of 'direct type-level causation' in macroeconomics.

2. Causal modeling in macroeconomics

When trying to justify macroeconomic policy decisions, macroeconomists typically employ causal models. A causal model can be defined as a triple $M = \langle V, U, \Pi \rangle$, where

(i) U is a set of background variables,

(ii) V is a set of causal structure variables,

(iii) Π is a set of parameters assigning a probability measure $P(u_i)$ to each U_i ($i = 1, \dots, n$) and a function

$$y_i = f_i(dc_i(y_i)),\ i = 1, \dots, n,$$

to each $Y_i \subseteq V$, where $DC(Y_i) \subseteq U \cup V \setminus Y_i$ is the set of variables representing direct causes of $Y_i \subseteq V$,

and where a direct cause is to be understood as causally relevant to (or a type-level cause of) and not as an actual (or token-level) cause of Y_i.[4]

The background variables in U encompass the influence of variables that represent direct causes of variables in V but have been omitted from V. The variables in V are measured variables that directly cause other variables in V, are directly caused by other variables in V or both. Like the variables in U and V, the parameters in Π are variables that represent sets of potential values that are measurable or quantifiable. When time-indexed, the variables in $U \cup V$ represent sets of ordered pairs that assign each possible value to each possible point in time. When Y_i is time-indexed, lagged values of Y_i may also show up among the direct causes of Y_i. Since the parameters assign a probability measure only to each of the variables in U, it is only

[4] Throughout the paper, uppercase letters will be reserved for variables and lowercase letters for their values. This convention is a bit unusual for (macro-) economists but widespread in the philosophical literature on causality.

the variables in U that supply a stochastic element. The underlying causal relationships between the variables in V are assumed to be deterministic.

Macroeconomists are interested in relations of direct type- and token-level causation alike: they are not only interested in a relation of direct type-level causation between, for instance, the real interest rate and aggregate output but also in relations of direct token-level causation between the real interest rate in, say, 2008Q4 and aggregate output in 2009Q2, between the real interest rate in 2016Q3 and aggregate output in 2019Q2 (where aggregate output in 2019Q2 is the future value that aggregate output will attain in the second quarter of 2019) etc. But they tend to understand direct token-level causes as instantiations or realizations of direct type-level causes, and not relations of direct type-level causation as generalizations of relations of direct token-level causation. The models that they use to justify macroeconomic policy decisions are therefore supposed to represent relations of direct type-level causation, and not relations of direct token-level causation.

As an example of a macroeconomic causal model, consider the rational expectations model that Hoover (2001, pp. 64-66; 2011, pp. 349-350; 2013, pp. 42-3) discusses at various points in his work. The model consists of two equations, of which the first describes the demand for real money:

(1) $m_t - p_t = \mu - \alpha({}_tp^e_{t+1} - p_t) + \nu_t$,

where subscripts index time, M_t is the nominal money stock, P_t the general price level, ${}_tP^e_{t+1}$ the expectation at time t of the price level at time t+1 so that $({}_tP^e_{t+1} - P_t)$ approximates the rate of inflation between t and t+1, μ a constant encompassing the influence of real GDP and the real interest rate[5], α a regression coefficient greater than zero and N_t an independent random error.[6] The second equation describes a monetary policy rule:

(2) $m_t = \lambda + m_{t-1} + \varepsilon_t$,

where λ stands for the constant growth rate of the money stock and E_t is an independent random error. The errors N_t and E_t are assumed to be serially uncorrelated, to be uncorrelated with each other and to have means of zero. If expectations of future inflation are formed rationally, i.e. if

(3) ${}_tp^e_{t+1} = E(p_{t+1} \mid \Xi_t)$,

where $E(\cdot)$ is the expectations operator and Ξ_t the information available at time t (which is assumed to include the structure of the economy and the current and lagged values of the

[5] Hoover (2001, p. 64) points out that the case in which the influence of real GDP and the real interest rate is encompassed by a constant is a case of hyperinflation, i.e. a case in which the inflation rate is so large relative to changes in the real interest rate and real incomes that it is reasonable to defer them as constants to the causal background. The case of hyperinflation isn't the norm but will be assumed here for purposes of simplification.

[6] In Hoover's model, M_t, P_t and ${}_tP^e_{t+1}$ are in natural logarithms. In macroeconomics, variables are usually lowercased when written in logarithms. In order to maintain the distinction between variables and their values, however, uppercase letters will be reserved for variables (no matter whether or not they are in logarithms) and lowercase letters for their values throughout the paper.

variables included in the model), then the inflation rate conforms to the rate of growth of the money stock:

(4) $\quad {}_{t}p^{e}_{t+1} - p_t = \lambda.$

Substituting (4) into (1) and rearranging yields:

(5) $\quad p_t = m_t - \mu + \alpha\lambda + \nu_t,$

which is the solution to the money-price model under rational expectations.

In macroeconomics, examples of causal models also include the dynamic-stochastic general-equilibrium (DSGE) models that central banks and other policymaking institutions use to justify macroeconomic policy decisions. Macroeconomic DSGE models are dynamic models built from microeconomic foundations with explicit assumptions about the behavior of the variables in U. As an example of a macroeconomic DSGE model, consider the following canonical new Keynesian model (cf. Romer, [4]2012, pp. 352-3):

(6) $\quad y_t = E_t[y_{t+1}] - r_t/\theta + u_t^{IS}$, $\qquad \theta > 0,$

(7) $\quad \pi_t = \beta E_t[\pi_{t+1}] + \kappa y_t + u_t^{\pi}$, $\qquad 0 < \beta < 1, \qquad \kappa > 0,$

(8) $\quad r_t = \varphi_{\pi} E_t[\pi_{t+1}] + \varphi_y E_t[y_{t+1}] + u_t^{MP} \qquad \varphi_{\pi} > 0, \qquad \varphi_y \geq 0,$

where Y_t (in logarithm) is aggregate demand, $E_t[Y_{t+1}]$ expectations in t of aggregate demand in t+1, R_t the real interest rate, Π_t (in logarithm) the rate of inflation, $E_t[\Pi_{t+1}]$ expectations in t of the rate of inflation in t+1. The U_t^{IS}, U_t^{π} and U_t^{MP} represent shocks to aggregate demand, inflation and the real interest rate, respectively, and are assumed to follow independent first-order autoregressive processes. The various parameters are identified in microeconomic theory: β and θ in the utility function of the representative household, κ in the price-setting behavior of the representative firm, and φ_{π} and φ_y in the (forward-looking) interest-rate rule followed by the central bank (cf. Romer, [4]2012, pp. 315-6, 329-31). While the model is obviously stylized (the dynamics of aggregate demand and inflation are very simple, the empirical performance of (7) is poor, everything is linear, all behavior is forward-looking etc.), it is “canonical” in the sense that it serves as a key reference point in macroeconomic DSGE modeling. Various modifications and extensions of it are used in central banks and other policymaking institutions.

As a final example of a macroeconomic causal model, consider the model that Angrist and Kuersteiner (2011, p. 727) use to analyze the causal effects of changes in the intended federal funds rate on changes in real GDP:

(9) $\quad \Delta gdp_{t+j}{}^{\psi}(\Delta ff_t).$

This expression describes changes in real GDP (ΔGDP_{t+j}) as a potential outcome: as the value that ΔGDP_{t+j} would assume if ΔFF_t (changes in the intended federal funds rate) were Δff_t. The model that Angrist and Kuersteiner think describes the process determining changes in the intended federal funds rate, i.e. the process that leads the members on the federal open markets committee (FOMC) to intend changes in the level of the federal funds rate, looks as follows:

$$\Delta ff_t = \psi(\Delta ff(z_t, t), \varepsilon_t, t),$$

where ψ is a general mapping and ε_t an independent regression error or monetary policy shock, i.e. an error reflecting the reaction of policymakers to idiosyncratic information. The set Z_t of observed random variables includes lagged, present and predicted values of variables standing for changes in real *GDP*, inflation and the unemployment rate.

The assumption that allows Angrist and Kuersteiner (2011, pp. 727-9) to understand ΔGDP_{t+j} as a potential outcome is the selection-on-observables assumption (SOA), i.e. the assumption that potential outcomes of ΔGDP_{t+j} are probabilistically independent of ΔFF_t, given Z_t.[7] Potential outcomes of changes in real *GDP* are probabilistically independent of ΔFF_t, given Z_t, if Z_t is an admissible (or de-confounding) set of covariates (for more on the property of admissibility cf. end of section 5 below). If Z_t is an admissible set of covariates, then a potential outcome of changes in real *GDP* can be identified from the product of the probability of changes in real *GDP*, given ΔFF_t and Z_t, and the probability of Z_t: $\sum_{Z_t} P(\Delta gdp_{t+j} | \Delta ff_t, z_t) \cdot P(z_t)$.

3. An interventionist account of macroeconomic causality

The decisive question is, of course, how membership in $DC(Y_i)$ is defined, i.e. what it means for a variable to be a direct type-level cause of Y_i. The present and the following two sections will consider three accounts that have been or can be proposed to answer that question. The main purpose of that consideration is to show that some of the conditions, that these accounts make reference to, are too restrictive in the case of macroeconomics. The first of the three accounts is Woodward's interventionist account. In the case where $i = 1$, it says that

(Ia) *X* directly type-level causes *Y* relative to *V* if and only if there is a possible intervention on *X* that changes *Y* (or its probability distribution) while all other variables in *V* remain fixed by intervention (cf. Woodward, 2003, pp. 55, 59).

What is striking about that definition is that it relativizes the notion of direct type-level causation to a variable set. According to Woodward (2003, p. 56), the selection of a variable set *V* depends on what an epistemic subject is prepared to accept as serious possibility. But what an epistemic subject is prepared to accept as serious possibility might be inter-subjectively different. Does Woodward believe that relations of direct type-level causation do not exist independently of the mind? No: he endorses a "kind of realism that […] is metaphysically modest and noncommittal"

[7] Angrist and Pischke (2009, p. 54) refer to the same assumption as "conditional independence assumption". In the present context, however, it is more appropriate to use the term "selection-on-observables assumption" because the assumption in question might otherwise be confused with the assumption that Angrist and Kuersteiner (2011, p. 729) refer to as "the key testable conditional independence assumption", i.e. with an assumption that involves actual, not potential outcomes.

(Woodward, 2003, p. 121). And this kind of realism can be reconciled with a de-relativized variant of (Ia) that says that

(Ib) *X* directly type-level causes *Y* if and only if there is a possible intervention on *X* that changes *Y* (or its probability distribution) while all other variables in- *and* outside *V* remain fixed by intervention.[8]

What it means to say that all other variables in- and outside *V* remain fixed by intervention becomes clear when Woodward's (2003, p. 98) definitions of the terms 'intervention' and 'intervention variable' are considered. His definition of 'intervention' says that *I*'s assuming some value z_i is an intervention on *X* with respect to *Y* if and only if *I* is an intervention variable for *X* with respect to *Y* and $I = z_i$ is an actual (token-level) cause of the value taken by *X*. Note that this definition is in line with Woodward's (2003, pp. 103-4) "nonanthropomorphism": *I*'s assuming some value z_i doesn't necessarily mean that *I* is set to z_i by policy or (more generally) human intervention.

Woodward's definition of the term 'intervention variable', by contrast, states that *I* is an intervention variable for *X* with respect to *Y* if and only if the following four conditions hold:

(*I*1) *I* type-level causes *X*,

(*I*2) certain values of *I* are such that when *I* attains these values, *X* is no longer determined by other variables that type-level cause it but only by *I*,

(*I*3) any directed path from *I* to *Y* goes through *X*, and

(*I*4) *I* is statistically independent of any variable *Z* that type-level causes *Y* and is on a directed path that does not go through *X*.

That all other variables in- and outside *V* remain fixed by intervention means that conditions (*I*1) – (*I*4) are satisfied: that there isn't any variable in- or outside *V* such that any of the propositions expressed by (*I*1) – (*I*4) becomes false. Satisfaction of conditions (*I*1) – (*I*4) guarantees the presence of a causal graph connecting *I* to *X* and *X* to *Y* by directed paths or the presence of a causal structure represented by that graph if there is a possible intervention on *X* that changes *Y* or its probability distribution.

Note that according to Woodward, an intervention on *X* that changes *Y* (or its probability distribution) while all other variables in- and outside *V* remain fixed only needs to be *possible* in order for a relation of direct type-level causation to hold between *X* and *Y*. Woodward is not entirely clear about the exact modality captured by the term 'possible' in 'possible intervention'. In one passage, he rejects the notion of an intervention on variables for which there is no well-defined notion of change (e.g. on variables standing for race, sex or species) (Woodward, 2003, pp. 113). In another passage, he admits interventions that even involve violations of physical laws (cf. Woodward, 2003, pp. 132). But what needs to be asserted at present is only that in

---

[8] For a more complete argument in favor of this de-relativized variant, cf. Henschen (2015: section 3).

order for X to directly type-level cause Y, it is not necessary that an intervention on an intervention variable satisfying (I1) – (I4) be actually present.

One might perhaps wonder why (Ia) and (Ib) make reference to "a" possible intervention on X that will change (the probability distribution of) Y when one holds fixed all other variables in- (and outside) V by intervention. Isn't a relation of direct type-level causation a lawful relation, and shouldn't (Ia) and (Ib) therefore say that X is a direct type-level cause of Y iff *all* interventions on X will change (the probability distribution of) Y_i when one holds fixed all other variables in- (and outside) V by intervention? Woodward (2003, pp. 68-70) answers this question in the negative. He points out that in order for a relation between X and Y to qualify as causal, this relation doesn't need to remain intact under all interventions on X. He cites examples from physics and the medical sciences to substantiate his point. But one might equally draw on examples from macroeconomics, including the models introduced in section 2 above.

Interventions, under which the relations represented by the equations of these models fail to remain intact, are of essentially four kinds. Interventions of the first kind set some of the variables of these equations to values that lie outside the domain of these equations. An example for an intervention of that kind is an intervention that sets M_t in equation (2) of the rational-expectation model from section 2 to zero. Interventions of the second kind result in a breakdown of the relation represented by an equation even though they set a variable that figures in that equation to a value that belongs to the domain of that equation. An example for an intervention of that kind is an intervention that in a neoclassical model, sets wages to values that are high enough to result in values of supplied labor that are sufficiently high to deprive a representative agent of the leisure she needs to recreate.

Interventions of the third kind lead to a breakdown of the relation represented by an equation by changing the value of a variable that is included in the set V of pre-selected variables but not included in that equation. An example for an intervention of that kind is an intervention that sets $E_t[\Pi_{t+1}]$ in equation (8) of the neo-Keynesian baseline model from section 2 to a value that yields a value for R_t that, given values for $E_t[Y_{t+1}]$ and U_t^{IS}, is sufficiently high to yield a negative value for Y_t in equation (6). Finally, interventions of the fourth kind lead to a breakdown of the relation represented by an equation by changing the value of a variable that is not even included in V. An example for an intervention of that kind is an intervention that changes the value of an omitted financial-market related variable in such a way that it counteracts the causal influence that Y_t is supposed to have on Π_t according to equation (7). Interventions of that kind are what Woodward refers to as "changes in the background circumstances" (cf. especially Hitchcock and Woodward, 2003, pp. 187-8). They are arguably the most interesting because they are the ones that typically get in the way of predictions of the behavior of aggregate quantities or of the consequences of macroeconomic policy manipulations.

Woodward holds that in cases where $i > 1$ (where we are dealing with systems of equations), the condition of modularity needs to be satisfied. Modularity requires that (i) each equation in the system satisfy the condition of invariance, and that (ii) for each equation, there be a possible intervention on the dependent variable that changes only that equation while the other equations in the system remain unchanged (cf. Woodward, 2003, p. 329). Clause (i) doesn't create any major difficulties: invariance just requires that an equation in which X figures as an independent variable and Y as a dependent variable remain invariant under some intervention on X so that Y will change under this intervention (cf. Woodward, 2003, p. 69). Invariance is of utmost importance for Woodward: he says that it "is the key feature a relationship must possess if it is to count as causal or explanatory" (Woodward, 2003, p. 237). Invariance also comes in degrees: the wider the range of possible interventions under which the equation expressing the relation of direct type-level causation remains invariant, the greater the degree of invariance of that equation (cf. Woodward, 2003, p. 257).

Clause (ii), by contrast, does create difficulties: Cartwright (2007, p. 15-6) shows that there are systems of equations that represent causal structures even though they do not satisfy that clause. Consider, first, an example of a system of equations that clearly satisfies that clause: the system of equations (6) – (8), i.e. the canonical new Keynesian model from section 2. An intervention on the real interest rate in the sense of Woodward is a manipulation of an intervention variable I that sets R_t in equation (8) to a particular value: that breaks all arrows that are directed into R_t and depart from any variables other than I (e.g. $E_t[\Pi_{t+1}]$ or $E_t[Y_{t+1}]$) in accordance with condition (I2). The modularity condition requires that setting R_t in equation (8) to a particular value doesn't disrupt any of the other equations in the system, i.e. equations (6) and (7). If the modularity condition is satisfied, condition (I3) is satisfied: the only path from I to Y_t will go through R_t. If, moreover, condition (I4) is satisfied, then R_t can be inferred to directly type-level cause Y_t.

Consider next an example of a system of equations that represents a causal structure even though it doesn't satisfy clause (ii): equations (2) and (5), i.e. Hoover's money-price model under rational expectations from section 2. Hoover (2013, p. 49) argues that an intervention in the sense of Woodward amounts to setting M_t to a particular value by breaking the causal arrow from M_{t-1} to M_t, and that breaking that arrow has the unwelcome consequence of rendering Λ meaningless, i.e. of undercutting any basis for forming rational expectations of the path of M_t. What is special about equations (2) and (5) is that they form a system that is subject to a nonlinear cross-equation restriction: Λ cannot be associated exclusively with either (2) or (5). It seems that Woodward's account doesn't hold for such systems: that a system of equations is non-modular whenever it is subject to a nonlinear cross-equation restriction.

One might object that nonlinear cross-equation restrictions result only if expectation variables are solved out, and that expectations variables don't need to be solved out because they can be observed: because macroeconomists can measure them by conducting surveys. Hoover (2001, p. 137) explains, however, why that objection is misguided:

> "True, people form expectations and act upon them […], but such expectations do not exist independently of the actions they affect; they are not palpable, like so many pounds of rice bought by a consumer […]. Of course, one could ask people to state their expectations. That, however, would be simply their guess about how they would act or would have acted in a situation that was not yet at hand or had already passed. Such expectations are no more directly observable than their own preferences and are subject to the same whimsy, arbitrariness, and adjustment to subtle changes in background conditions."

Hoover suggests that expectations fall into the same category as preferences. In revealed-preference theory, a consumer's preference is reconstructed from her behavior (from her "revealed" preference). Statements about what she thinks she prefers are to be dismissed as neither verifiable nor trustworthy.[9] Similarly, expectation variables cannot be measured because a subject's statement about what she expects can neither be verified nor trusted. Expectation variables therefore need to be solved out. If they are not solved out (as in the case of the canonical new Keynesian model), they need to be interpreted as attaining whatever value is required to render the model consistent.

It is important to see, however, that Woodward's account can be modified in order to accommodate systems of equations that are subject to nonlinear cross-equation restrictions. In order to see this, note that such systems violate clause (ii) because they violate condition (*I*2). Equations (2) and (5) do not violate clause (ii) in the sense that changing equation (2) disrupts equation (5): setting M_t to m_t while breaking the causal arrows from Λ and M_{t-1} to M_t renders Λ "meaningless" but doesn't disrupt equation (5). Equations (2) and (5) violate clause (ii) because for equation (2), there isn't any possible intervention on M_t that changes that equation: because for equation (2), there isn't any possible intervention that satisfies condition (*I*2), i.e. that breaks all arrows directed into M_t and departing from Λ and M_{t-1}.

If systems of equations that are subject to nonlinear cross-equation restrictions violate clause (ii) because they violate condition (*I*2), then dropping condition (*I*2) from the conditions that Woodward lists to define 'intervention variable' will modify his account in such a way that it accommodates such systems. It will modify his account in essentially two ways. It will, first, lead to a reformulation of clause (ii) of his modularity condition: to a formulation requiring that for each equation, there be a possible intervention on the dependent variable that satisfies (*I*1), (*I*3) and (*I*4) and does not disrupt any of the other equations in the system. Cartwright's example of a

[9] Sen (1973: 242) provides a long list of quotations that testify to the worries about introspection and verifiability that motivated revealed-preference theory.

carburetor is one that violates even this reformulation. But in macroeconomics, systems of equations generally satisfy this reformulation or (as was shown earlier) even clause (ii). They even always satisfy this reformulation or clause (ii) if the notion of an intervention is not restricted to either interventions on parameters (variables in Π) or interventions on variables in V (as I shall argue in the following section).

The second modification relates to the phrase "while all other variables … remain fixed by intervention" in definitions (Ia) and (Ib). Definitions (Ia) and (Ib) would no longer define '*direct* type-level causation' unless interventions that hold fixed all other variables were understood as manipulations of intervention variables that satisfy conditions (I1) – (I4) including (I2). In order to see this, consider the case of X causing Y only indirectly via W, i.e. the case of a directed path from X to Y through W. We would not be able to rule out that case unless there was an intervention variable certain values of which are such that when it attains these values, W is no longer determined by other variables that type-level cause it but only by that intervention variable. X might, after all, be among those other variables.

But once we start tampering with the phrase the phrase "while all other variables …", we should also note that in order for definitions (Ia) and (Ib) to define 'direct type-level causation', it suffices to hold fixed only the causal parents of Y excluding X, i.e. all variables in $DC(Y) \setminus X$. Woodward (2003, p. 55) suggests that the phrase "while all other variables …" is needed to ensure that the causal relation between X and Y is direct. But Pearl ([2]2009, pp. 127-8) is right when pointing out that in order to ensure that the causal relation between X and Y is direct, it suffices to hold constant the parents of Y excluding X. The phrase "while all other variables … remain fixed by intervention" should therefore be restated as "while all causal parents of Y excluding X … remain fixed by intervention", where an intervention that holds fixed any of the variables in $DC(Y) \setminus X$ is to be understood as different from an intervention on X. I am going to spell out that difference in greater detail in section 6.

4. Macroeconomic causality as privileged parameterization

The second of the three accounts that have been or can be proposed to define membership in $DC(Y_i)$ is Hoover's account. What is the definition that is central to that account? That question is a bit difficult to answer because Hoover nowhere states any definition that is as explicit as Woodward's, and because he in fact offers two definitions. A condition that figures in both definitions requires that in a self-contained (or soluble) system of equations, X and Y belong to different self-contained subsystems, that the self-contained subsystem including X be of a lower order than the self-contained subsystem including Y (i.e. that the former needs to be solved in order to solve the second), and that there be no self-contained subsystem intervening between the two. Hoover (2012, p. 782) refers to this condition as "Simon's hierarchy condition" in order to emphasize the indebtedness of his approach to the work of Herbert Simon.

In Hoover's first definition, Simon's hierarchy condition combines with the condition that the parameterization of the self-contained system of equations (i.e. that the set Π of parameters in a causal model *M*) be privileged, where a parameterization is privileged if it (and the functional form of the causal model equations) is invariant "in the face of specific interventions", i.e. in the face of interventions on the parameters that we can directly control to control for the variables in *V* indirectly, and where a parameterization is invariant in the face of such interventions if the parameters in Π are variation-free, i.e. if they are mutually unconstrained. Hoover (2011, pp. 344-5; 2012, pp. 781-2; 2013, pp. 40-1) points out that Simon introduces this condition to solve the problem of the observational equivalence of different self-contained systems of equations. This condition may therefore be referred to as "Simon's condition of privileged parameterization".

In Hoover's second definition, by contrast, Simon's hierarchy condition combines with Hoover's "parameter-nesting condition" (cf. Hoover, 2012, p. 782), i.e. with the condition that $\Pi(X)$ be a proper subset of $\Pi(Y)$, where $\Pi(X)$ and $\Pi(Y)$ are the sets of parameters that can be controlled directly to control *X* and *Y* indirectly. Hoover (2012, p. 782) suggests that the parameter-nesting condition extends Simon's approach, while also inheriting the property that causal order is uniquely defined by the functional relations among variables, as long as we know the privileged parameters.

Hoover's first definition can be reconstructed from what he says about privileged parameterizations (cf. Hoover 2011, p. 345; 2013, p. 41). It states that

(IIa) *X* directly type-level causes *Y* if and only if Simon's hierarchy condition and Simon's condition of privileged parameterization hold.

Hoover's second definition, by contrast, can be extracted from his set-theoretic formalization of causal structure (cf. Hoover, 2001, pp. 61-3). It says that

(IIb) *X* directly type-level causes *Y* if and only if Simon's hierarchy condition and Hoover's parameter-nesting condition hold.

Which definition should we prefer? It seems to me that (IIa) is preferable. In one passage, Hoover (2012, p. 782) suggests that (IIb) is required if systems of equations that are subject to nonlinear cross-equation restrictions are to be accommodated. But in the same passage, he shows that the parameter-nesting condition is not necessary: that the equations $A = \alpha_A$ and $B = \alpha_A A$ might represent a causal structure in which *A* directly type-level causes *B* even though $\Pi(A) = \Pi(B) = \{\alpha_A\}$, i.e. even though $\Pi(A)$ is not a proper subset of $\Pi(B)$. There are, moreover, at least two passages in which Hoover (2011, p. 350; 2013, p. 42) claims that (IIa) can successfully deal with the money-price model under rational expectations from section 2.[10] And I agree. In the case of that model, Simon's hierarchy condition and his condition of privileged parameterization

[10] In these passages, the terms "Simon's framework" and "structural account" must refer to (IIa) because neither article is concerned with the parameter-nesting condition.

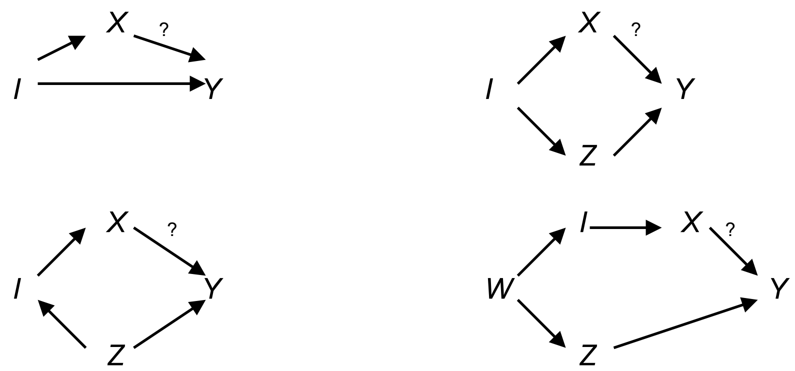
hold: the parameters of that model are variation-free; and equations (2) and (5) form a self-contained system in which (2) is of lower order and no subsystem intervenes.

There are, however, at least two problems with Hoover's characterization of the notion of intervention, as it is basic to both (IIa) and (IIb). The first problem is that Hoover (2001, p. 61; 2011, pp. 346-7; 2013, pp. 52-3) restricts that notion to interventions on parameters (or variables in Π). He shares Woodward's non-anthropomorphism: he doesn't think that direct parameter control is necessarily control through policy manipulation; he accepts e.g. war-induced changes of parameter values as cases of direct parameter control (cf. Hoover 2001, pp. 229-230). Hoover (2011, p. 348) is also prepared to understand variation-free parameters as intervention variables. But an intervention for Hoover is not the same as for Woodward: while for Hoover it changes the value of a variable in Π, for Woodward it changes the value of a variable in V.

The problem with this restriction is that in macroeconomics, there are obviously interventions that cannot be understood as interventions on parameters. Take the canonical new Keynesian model from section 2: how does the Federal Reserve intervene on R_t in equation (6)? It changes R_t (i.e. the real interest rate which is simply the nominal interest rate minus the rate of inflation) by changing the nominal interest rate and the nominal interest rate by supplying or removing reserves from the banking system. A change of that kind is not an intervention on any of the parameters in the model. Instead, it is an intervention on R_t in the sense of Woodward. Woodward's characterization of the notion of intervention is, of course, likewise too restrictive. If Hoover's money-price model under rational expectations is taken as a point of departure, it is legitimate to understand a monetary policy intervention as an intervention on Λ. The general lesson to be drawn states that in macroeconomics, the notion of an intervention is not to be restricted to either interventions on parameters (variables in Π) or interventions on variables in V.

The second problem with Hoover's characterization is his insistence that in order for X to directly type-level cause Y, there doesn't need to be a possible intervention on a parameter-intervention variable that satisfies all of Woodward's conditions (I1) – (I4) (cf. Hoover, 2011, pp. 348-9). Hoover agrees that in order for X to directly type-level cause Y, there needs to be a possible intervention on a parameter-intervention variable that satisfies (I1). But he denies that there needs to be a possible intervention on a parameter-intervention variable that satisfies any of (I2) – (I4). And while this denial is justified in the case of (I2), it is unjustified in the cases of (I3) and (I4). It is justified in the case of (I2) because, as we have seen toward the end of section 3, there are macroeconomic systems of equations, for which there is no possible intervention satisfying (I2), i.e. systems of equations that are subject to nonlinear cross-equation restrictions. But it is unjustified in the cases of (I3) and (I4) because a parameter-intervention variable is causally relevant to X if it satisfies (I1), because a parameter-intervention variable corresponds to an arrow-emitting node if it is causally relevant, and because *X* might not qualify as a direct type-level cause of *Y* if *Y* is connected to the parameter-intervention variable by a path that doesn't go through *X*.

Conditions (*I*3) and (*I*4) are, of course, meant to rule out the existence of such a path, as becomes clear in the figure that Woodward (2003, pp. 101-2) uses to illustrate conditions (*I*3) and (*I*4):



Hoover insists that it is impermissible "to draw causal arrows from parameters to other parameters and variables" (Hoover, 2012, p. 781), and that drawing such arrows is against his "representational conventions" (Hoover, 2010, p. 391). Hoover refers to these conventions when discussing an example that Cartwright (2007, p. 207) thinks is a counterexample to his parameter-nesting condition. I'm not sure whether these conventions make sense: if we can control a parameter directly to control a variable indirectly, then that parameter is causally relevant to (or a type-level cause of) that variable (Hoover admits that much when agreeing that an intervention on a parameter-intervention variables satisfies Woodward's condition (*I*1)); and if that parameter is causally relevant to that variable, then why shouldn't we be allowed to draw a causal arrow from that parameter to that variable (i.e. to understand that parameter as corresponding to an arrow-emitting node in a causal graph)? I would also like to emphasize, however, that the above argument does not rely on a violation of Hoover's conventions. When avoiding the language of causal graphs it says that on Hoover's account, an intervention on a parameter-intervention variable needs to satisfy (*I*3) and (*I*4) because that parameter-intervention variable is causally relevant if it satisfies (*I*1), and because *X* might not qualify as a direct type-level cause of *Y* if the parameter-intervention variable type-level causes or is type-level caused by a variable (or set of variables) *Z* that type-level causes *Y*.

If an intervention on a parameter-intervention variable needs to satisfy (*I*1), (*I*3) and (*I*4) in order for *X* to directly type-level cause *Y*, then an important question relates to the exact relationship between Hoover's account and the macroeconomic variant of Woodward's account. In the present paper, I can only offer a conjecture (instead of a formal proof). Both accounts consist of essentially three components: one component guaranteeing that the relation of type-level causation between *X* and *Y* is direct (the reformulation of the phrase "while all other variables …"

and Simon's hierarchy condition), one component specifying modularity (the reformulation of Woodward's modularity condition and Simon's condition of privileged parameterization), and one component characterizing the notion of an intervention (Woodward's definitions of 'intervention' and 'intervention variable' in terms of (*I*1), (*I*3) and (*I*4) and Hoover's characterization of direct parameter control that he mistakenly believes can get along without conditions (*I*3) and (*I*4)). The conjecture is that both accounts are equivalent, once the notion of an intervention is not restricted to either interventions on parameters (variables in Π) or interventions on variables in V.

I sufficiently motivated that conjecture with respect to the third component. It is also sufficiently motivated with respect to the first component: while Simon's hierarchy condition is meant to guarantee direct causation in the case of systems of equations, Woodward's clause "while all other variables … " is more general and guarantees direct causation also in more garden-variety cases. How about the modularity-specifying component? Hoover (2012, p. 782) agrees that Simon's condition of privileged parameterization is a modularity condition. He says, more specifically, that the condition that parameters be variation-free is a modularity condition. But the condition that parameters be variation-free is a precondition for their being privileged. Simon's condition of privileged parameterization may therefore be understood as a modularity condition. Hoover (2012, p. 782) also suggests that his (or Simon's) and Woodward's modularity conditions are different. But while Hoover's (or Simon's) modularity condition says that parameters are unconstrained, Woodward's modularity condition says that functions (or the equations expressing them) are unconstrained. And remember from the definition of the notion of a causal model in section 2 that the parameters of that model assign these functions to each of the Y_i. So how can Woodward's and Hoover's modularity conditions be different?

5. The potential outcome approach to macroeconomic causality

The third definition of membership in $DC(Y_i)$ can be extracted from the potential-outcome approach that Angrist and Kuersteiner (2011) have introduced into macroeconomics more recently. It says that

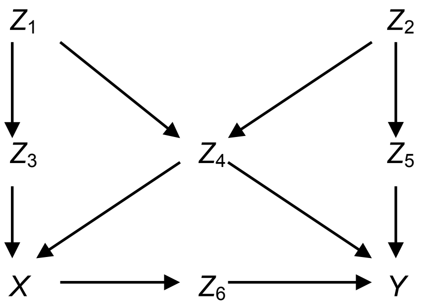
(III) X is a direct type-level cause of Y if and only if

- (*a*) $\sum_Z P(y \mid x, z) \cdot P(z) \neq 0$,
- (*b*) potential outcomes of Y are probabilistically independent of X given Z, and
- (*c*) all variables in $DC(Y) \setminus X$ remain fixed by intervention,

where the expression in (*a*) measures the causal effect of X on Y, where (*b*) is the selection-on-observables assumption (or SOA), and where (*c*) is meant to ensure that (III) defines the notion of *direct* type-level causation. It is to be conceded that Angrist and Kuersteiner's approach contains only conditions (*a*) and (*b*), and that conditions (*a*) and (*b*) combine to define the notion of total cause, not that of direct type-level cause. But the purpose of the present section is to consider whether Angrist and Kuersteiner's approach might pass for a valid account of direct type-level causation in macroeconomics. And that consideration requires that their account be translated into an account of direct type-level causation. Fortunately, condition (*c*) is exactly what is needed to achieve that translation (cf. end of section 3).

In his discussion of the literature on potential outcomes, Pearl (^{2}2009, p. 80) refers to the conditioning set Z of variables in the SOA as "admissible" or "deconfounding". He argues that researchers need a workable criterion to guide their choice of the covariates that need to be included in that set. And he claims that the concept of the SOA "falls short of providing researchers with a workable criterion to guide the choice of covariates" (Pearl, ^{2}2009, p. 79). Pearl's claim is put forth as a general criticism of the potential outcome framework. But it clearly applies to the monetary-policy model from section 2: all that Angrist and Kuersteiner (2011, p. 729) remark with respect to the SOA that they think holds for Z_t, ΔFF_t and potential outcomes of ΔGDP_{t+j} is that it holds "after appropriate conditioning"; and all they do to justify their choice of covariates is to conduct a series of diagnostic (or misspecification) tests on a number of causal models for the process determining ΔFF_t.

The criterion that Pearl (^{2}2009, p. 79) himself provides is the "back-door criterion". He says that Z satisfies the backdoor criterion relative to an ordered pair of variables (X, Y) in a causal graph if Z (i) doesn't include any descendants of X and (ii) blocks every path between X and Y that contains an arrow into X, where Z is said to "block" a path p if p contains at least one arrow-emitting node that is in Z or at least one collision note that is not in Z and has no descendants in Z (cf. Pearl, ^{2}2009, pp. 16-7). Pearl (^{2}2009, p. 80) uses the following graph to illustrate the backdoor criterion:



In this graph, $\{Z_3, Z_4\}$ and $\{Z_4, Z_5\}$ satisfy the backdoor criterion because they do not include any descendants of X, and because they block every path between X and Y that contains an arrow into X. It is clear that $\{Z_6\}$ does not satisfy the backdoor criterion because it is a descendent of X. Note, however, that $\{Z_4\}$ by itself doesn't satisfy the back-door criterion either: it blocks the path $X \leftarrow Z_3 \leftarrow Z_1 \rightarrow Z_4 \rightarrow Y$ because the arrow-emitting node is in Z; but it does not block the path $X \leftarrow Z_3 \leftarrow Z_1 \rightarrow Z_4 \leftarrow Z_2 \rightarrow Z_5 \rightarrow Y$ because none of the arrow-emitting nodes (Z_1, Z_2) is in Z, and because the collision node Z_4 is not outside Z. Pearl (^{2}2009, pp. 80-81) then proves the proposition that $P(y|do(x)) = \sum_Z P(y|x, z) \cdot P(z)$ if and only if Z satisfies the backdoor criterion, where $P(y|do(x))$ is the probability that $Y = y$ if X is set to x by intervention.

For the case of Angrist and Kuersteiner's monetary-policy model from section 2, all this means that ΔFF_t (changes in the intended federal funds rate) cannot directly type-level cause ΔGDP_{t+j} (changes in real GDP) unless $\sum_{Z_t} P(\Delta gdp_{t+j}|\Delta ff_t, z_t) \cdot P(z_t) \neq 0$ and ΔGDP_{t+j} are probabilistically independent of ΔFF_t given Z_t; that ΔGDP_{t+j} cannot be probabilistically independent of ΔFF_t given Z_t unless Z_t is admissible; and that Z_t cannot be admissible unless the variables included in Z_t (lagged, present and predicted values of ΔGDP_{t+j}, inflation and the unemployment rate) satisfy the back-door criterion: unless Z_t doesn't include any descendants of ΔFF_t and blocks every path between ΔFF_t and ΔGDP_{t+j} that contains an arrow into ΔFF_t.

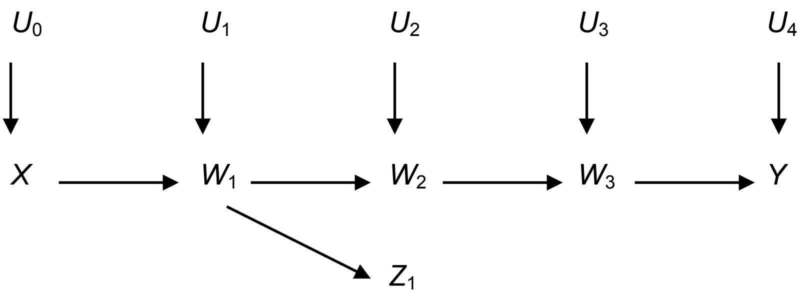
It is true that Angrist and Kuersteiner do not explicitly endorse the backdoor criterion as a criterion to guide their choice of the variables in Z_t. But Pearl *proves* that $\sum_Z P(y|x, z) \cdot P(z)$ cannot measure the causal effect of X on Y unless the variables in Z satisfy the backdoor criterion. It is accordingly fair to say that (III) cannot qualify as an adequate definition of 'direct type-level causation' unless the variables included in Z satisfy the backdoor criterion. Since Angrist and Kuersteiner do not explicitly endorse the backdoor criterion and condition (*c*) in (III) does not figure explicitly in their approach, it is questionable whether (III) can be "extracted" from their approach. But the approach to which (III) is central should be given a name. And since Angrist and Kuersteiner (2011) are the ones to have introduced the potential outcome approach into macroeconomics, I will continue to refer to (III) as central to Angrist and Kuersteiner's approach.

It is important to understand that the backdoor criterion implies Woodward's conditions (*I*1) – (*I*4). For Pearl (^{2}2009, pp. 70-71), *do*(*x*) amounts to setting X to x by manipulating an intervention variable I and by breaking all arrows directed into X and departing from variables other than I. For Pearl, that is, *do*(*x*) requires that Woodward's conditions (*I*1) and (*I*2) be satisfied. Condition (ii) of the backdoor criterion, moreover, rules out the same cases as Woodward's conditions (*I*3) and (*I*4). The cases that Woodward's conditions (*I*3) and (*I*4) are meant to rule out correspond to the four graphs in figure 1 (cf. end of section 4). These are cases in which it is impossible to infer that X directly type-level causes Y if Z is unknown. These are also cases in which Z blocks the paths between X and Y that contain an arrow into X if it is known. Therefore, conditioning on Z (knowing the value of Z) rules out the same cases as conditions (*I*3) and (*I*4). The backdoor criterion is a bit stronger than Woodward's conditions (*I*1) – (*I*4) since condition (i) of the backdoor criterion rules out all cases in which arrows are directed into Z and depart from X. But the backdoor criterion represents a set of conditions that includes Woodward's conditions (*I*1) – (*I*4). One may accordingly say that the backdoor criterion implies these conditions.

If the backdoor criterion implies Woodward's conditions (*I*1) – (*I*4), then an important question relates to the exact relationship between the macroeconomic variant of Woodward's interventionist account and Angrist and Kuersteiner's potential-outcome approach. As in the case of the relation between that variant and Hoover's account, I can only offer a conjecture. Angrist and Kuersteiner's potential-outcome approach consists of essentially two components: one component guaranteeing that the relation of type-level causation between X and Y is direct (condition (*c*) in definition (III)), and one component characterizing the notion of an intervention (conditions (*a*) and (*b*) together with the equality of $P(y \mid do(x))$ and $\sum_Z P(y \mid x, z)\cdot P(z)$, given that Z satisfies the backdoor criterion). A component specifying modularity is missing because (9) expresses the only functional relationship considered. But such a component could easily be added by requiring e.g. like Woodward that setting X to x by intervention do not disrupt any other functional relationships.

The conjecture is that the potential-outcome approach will reduce to the macroeconomic variant of Woodward's account, if two conditions are dropped: condition (i) of the backdoor criterion and Woodward's condition (*I*2). The loss of (*I*2) can be compensated if an intervention that holds fixed a variable in $DC(Y) \setminus X$ is interpreted as a manipulation of an intervention variable for $W \in DC(Y) \setminus X$ with respect to Y that breaks all arrows directed into W and departing from variables other than that intervention variable. And the loss of condition (i) of the backdoor criterion can be compensated if acyclicity is accepted as a general property of the causal graphs that can be taken to represent relations of direct type-level causation in macroeconomics.

Pearl (^{2}2009, p. 339) points out that estimates of $\sum_Z P(y \mid x, z)\cdot P(z)$ may be biased if the variables in Z do not satisfy condition (i) of the backdoor criterion. He uses the following graph to illustrate that point:



This graph is supposed to be a fine-grained variant of $X \rightarrow W_1 \rightarrow W_2 \rightarrow W_3 \rightarrow Y$. If Z_1 is included in Z, then estimating $\sum_Z P(y \mid x, z)\cdot P(z)$ will be biased because W_1 is a collider, i.e. because there will be a dependency between X and U_1 that in the coarse-grained variant of the graph, acts like a backdoor path: $X \leftrightarrow U_1 \rightarrow W_1 \rightarrow W_2 \rightarrow W_3 \rightarrow Y$. Pearl (^{2}2009, p. 339) suggests that condition (i) of the backdoor criterion is meant to rule out such backdoor paths. But he also admits that the presence of descendants of X does not necessarily bias estimates of $\sum_Z P(y \mid x, z)\cdot P(z)$: that e.g. a descendant of X that is not also a descendant of some W_i or U_i (with $i > 0$) can safely be conditioned on without introducing bias.

If the presence of descendants of X does not necessarily bias estimates of $\sum_Z P(y|x, z)\cdot P(z)$, then condition (i) of the backdoor criterion is unnecessarily strong. That condition becomes even superfluous once acyclicity is accepted as a general property of the causal graphs that can be taken to represent relations of direct type-level causation. If acyclicity is accepted as a general property, then a causal graph like $X \leftrightarrow U_1 \rightarrow \ldots$ is inadmissible from the start. In macroeconomics, the causal graphs that can be taken to represent causal models are typically acyclic. Simultaneous equation models require representation by cyclical causal graphs (cf. Pearl, ^{2}2009, pp. 28, 215, for an example from economics) but do not qualify as causal in the sense of the definition stated at the beginning of section 2. Condition (i) of the backdoor criterion is therefore not a necessary condition for direct type-level causation in macroeconomics.

6. An adequate account of macroeconomic causality

The analysis of the preceding three sections suggests that a macroeconomic variant of Woodward's interventionist account qualifies as an adequate account of macroeconomic causality: that in macroeconomics, an adequate definition of 'direct type-level causation' says that

(*) X directly type-level causes Y if and only if there is a possible intervention on X that changes Y (or its probability distribution) while all causal parents of Y except X remain fixed by intervention,

where

- I_X's assuming some value z_i is an intervention on X with respect to Y if and only if I_X is an intervention variable for X with respect to Y and $I_X = z_i$ is an actual (token-level) cause of the value taken by X.
- I_X is an intervention variable for X with respect to Y if and only if the following three conditions hold:
  - (I_X1) I_X type-level causes X,
  - (I_X3) any directed path from I_X to Y goes through X, and
  - (I_X4) I_X is statistically independent of any variable Z that type-level causes Y and is on a directed path that does not go through X.
- I_W's assuming some value z_i is an intervention on $W \in DC(Y) \setminus X$ with respect to Y if and only if I_W is an intervention variable for W with respect to Y and $I_W = z_i$ is an actual (token-level) cause of the value taken by W.
- I_W is an intervention variable for W with respect to Y if and only if the following four conditions hold:
  - (I_W1) I_W type-level causes X,
  - (I_W2) certain values of I_W are such that when I_W attains these values, W is no longer determined by other variables that type-level cause it but only by I_W,

(I_W3) any directed path from I_W to Y goes through W, and

(I_W4) I_W is statistically independent of any variable Z that type-level causes Y and is on a directed path that does not go through W.

- an intervention variable is either a parameter (i.e. a variable in Π) or a variable in V.

It is easy to see that this definition coincides with definition (Ia) or (Ib), except that "all other variables …" is replaced with "all causal parents of Y except X", that interventions on X and interventions on variables in $DC(Y) \setminus X$ are interpreted differently, and that the notion of an intervention is no longer restricted to interventions on variables in V.

I argued in section 3 that interventions on X and interventions on variables in $DC(Y) \setminus X$ need to be interpreted differently because Woodward's account cannot accommodate cases of macroeconomic models that are subject to nonlinear cross-equation restrictions unless (I2) is dropped from the conditions that Woodward lists to define the term 'intervention variable for X with respect to Y'. I also argued in section 3 that dropping condition (I2) has the unwelcome consequence of rendering the relation of type-level causation between X and Y indirect unless Woodward's phrase "while all other variables …" is replaced with "while all causal parents of Y except X …", and unless interventions that hold fixed all causal parents of Y except X are interpreted as manipulations of intervention variables that satisfy conditions (I_W1) – (I_W4). I pointed out in section 4 that allowing intervention variables to include parameters is necessary because in macroeconomics, it is sometimes legitimate to understand interventions as interventions on parameters.

Definition (*) is an adequate definition of 'direct type-level causation' in macroeconomics. I argued toward the end of the preceding section that it doesn't coincide with the definition that can be extracted from the potential-outcome approach that Angrist and Kuersteiner (2011) have introduced into macroeconomics more recently: that the latter is inadequate as a definition of 'direct type-level causation' in macroeconomics, and that it reduces to the former if Woodward's condition (I2) and condition (i) of the backdoor criterion are dropped. But I also conjectured toward the end of section 4 that definition (*) coincides with Hoover's definition (IIa), as long as the notion of intervention that is basic to Hoover's definition is not restricted to interventions on parameters (or variables in Π).

By way of conclusion, I'd like to point to a common feature of all three definitions: their circularity. Causal vocabulary ('cause', 'intervention', 'control', 'arrow') shows up on both sides of the biconditionals expressed by (IIa), (III) and (*). While this is obvious in the case of (*), it is less obvious in the cases of (IIa) and (III). Hoover (2013, p. 52) claims at one point that "the structural account does not actually *use* the notion of direct control to define causal order". But that claim is true only as long as Simon's condition of privileged parameterization is not spelled out. Hoover is aware, of course, that spelling out that condition requires the notion of direct control. In a different context, he even says that "[i]nvariance of the functional forms in the face of specific interventions is […] the hallmark of a true causal representation" (cf. Hoover, 2011, p. 345). One might likewise claim that the potential-outcome approach doesn't actually use any causal vocabulary to define the notion of direct type-level causation. But similarly, this claim is true only as long as the backdoor criterion isn't spelled out as a condition that guides the choice of covariates.

Definitional circularity is arguably a problem for philosophers who aim to develop reductive accounts of causality: theories that Glymour (2004, pp. 779-80) refers to as "Socratic" (not without a sense of irony, of course), and that he distinguishes from Euclidean theories. Woodward (2003, pp. 104-5), however, makes it clear that he aims to develop a Euclidean, and not a Socratic theory: that he is interested in the conceptual entanglement between causation and intervention, and not in any non-circular definition or reductive account of direct type-level causality. [11] And similarly, Hoover (2001, p. 42) claims that "[circularity] is less troubling epistemologically than it might seem to be ontologically".

References:

Angrist, J. D. and Kuersteiner, G. M. (2011). Causal effects of monetary shocks: semi-parametric conditional independence tests with a multinomial propensity score. *The Review of Economics and Statistics* 93(3), 725-747.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton, NJ: PUP.

Cartwright, N. (2007). *Hunting causes and using them*. Cambridge: CUP.

Cartwright, N. (2009). Causality, Invariance, and Policy. In Kincaid, H. and Ross, D. (eds.), *The Oxford Handbook of the Philosophy of Economics*. Oxford: OUP, 410-423.

Engle, R. F., Hendry, D. F. and Richard, J. F. (1983). Exogeneity. *Econometrica* 51(2), 277-304.

Glymour, C. (2004). Critical Note on: James Woodward, Making Things Happen. *British Journal for the Philosophy of Science* 55, 779-790.

Granger, C. W. J. (1980). Testing for Causality: A Personal Viewpoint. *Journal of Economic Dynamics and Control* 2(4), 329-352.

Henschen, T. (2015). *Ceteris paribus* laws and the interventionist account of causality. *Synthese* 192(10), 3297-3311.

Hitchcock, C. and Woodward, J. (2003). Explanatory Generalizations, Part II: Plumbing Explanatory Depth. *Nous* 37(2), 181-199.

Hoover, K. D. (2001). *Causality in Macroeconomics*, Cambridge: CUP.

Hoover, K. D. (2004). Lost Causes. *Journal of the History of Economic Thought* 26(2), 149-164.

[11] I should mention, however, that Glymour (2004: 785, 790) also argues that the notion of a causal graph is basic, and that the notion of an intervention should be defined in terms of it.

Hoover, K. D. (2010). Causal Pluralism and the Limits of Causal Analysis: A Review of Nancy Cartwright's *Hunting Causes and Using Them*. In Emmett, R. B. and Biddle, J. E. (eds.), *Research in the History of Economic Thought and Methodology*, Vol. 28A, 381-395.

Hoover, K. D. (2011). Counterfactuals and Causal Structure. In McKay Illari, P., Russo, F. and Williamson, J. (eds.), *Causality in the Sciences*. Oxford: OUP, 338-360.

Hoover, K. D. (2012). Causal Structure and Hierarchies of Models. *Studies in History and Philosophy of Science Part C* 43(4): 778-86.

Hoover, K. D. (2013). Identity, Structure, and Causal Representation in Scientific Models. In Chao, H.-K., Chen, S.-T. and Millstein, R. (eds.). *Towards the Methodological Turn in the Philosophy of Science: Mechanism and Causality in Biology and Economics*. Dordrecht: Spinger, 35-60.

Maddala, G. S. and Lahiri, K. ([4]2009). *Introduction to Econometrics*. Chichester: Wiley & Sons.

Pearl, J. ([2]2009). *Causality: Models, Reasoning, and Inference*. Cambridge, MA: Cambridge University Press.

Romer, D. ([4]2012). *Advanced Macroeconomics*. New York: McGraw-Hill.

Sen, A. (1973). Behavior and the Concept of Preference. *Economica* 40(159), 241-259.

Spirtes, P., Glymour, C., Scheines, R. (1993). *Causation, Prediction and Search*, New York: Springer.

Woodward, J. (2003). *Making Things Happen: a Causal Theory of Explanation*. Oxford: OUP.