Economics and Business Review, Vol. 3 (17), No. 3, 2017: 80-99
DOI: 10.18559/ebr.2017.3.5

Perspectival representation in DSGE models[1]

Paweł Kawalec[2]

Abstract: DSGE models (Introduction) have recently been criticized by P. Romer (2016) as pseudoscientific (Section 1). Their dominance is attributed to the uncritical "deference to authority" that has dominated macroeconomics "for the last 30 years". In contrast, the paper aims to support the widespread view that – their problems notwithstanding – DSGE models meet the epistemic standards of scientific research. The argument turns on the recent advancements in theories of scientific representation (Section 1) and of empirical grounding (Section 2). The latter is illustrated with a historical case, which also substantiates Romer's constructive point on the role of theory in design of measurements.

Keywords: DSGE models, scientific representation, objectivity, empirical grounding.

JEL codes: B41, C11, E17.

Introduction: An outline of DSGE models

The presentation and discussion of Romer's criticisms which concern epistemic norms of scientific research,[3] is preceded by an outline of the methodological characterization of DSGE models (Grabek, Kłos, & Koloch, 2010; Christiano, Trabandt, & Walentin, 2010; Del Negro & Schorfheide, 2013; Lindé, Smets, & Wouters, 2016; Breuss, 2016; Nachane, 2016). They gradually emerged as an outcome of progressive transformations within macroeconomics which were initiated by criticisms of the approach elaborated by the Cowles Commission. The transformations are commonly recognized as a clear advancement, per-

[1] Article received 17 May 2017, accepted 5 August 2017.
The author gratefully acknowledges the support of the Polish National Science Center (NCN) under the grant no UMO-2014/15/B/HS1/03770. I appreciate the comments of the Reviewers, which allowed me to improve the argument presented in the paper and also to incorporate the bifurcation theory.

[2] John Paul II Catholic University of Lublin, Faculty of Philosophy, Institute of Theoretical Philosophy, Al. Racławickie 14, 20-950 Lublin, Poland, pawel.kawalec@kul.pl.

[3] Romer's paper also takes up substantial debates in macroeconomics, e.g. the role of monetary factors for aggregate economic activity; for a response see (Williamson, 2017).

haps even a paradigm-shift: "The transformation of macroeconomics that Lucas initiated had all the trappings of a Kuhnian scientific revolution: a shift in the type of issues addressed, a new conceptual toolbox, new mathematical methods, the rise to power of a new generation of scholars" (Vroey, 2016, p. 151).[4]

The outline to follow is intended to enrich our understanding of the heterogeneous elements of DSGE models by displaying their progressive development, while also pointing out relevant analogies with the historical case to be discussed in Section 3.[5] In the 1940's many governments initiated systematic statistical reporting of national economic activities and that created the opportunity for economists to construct and test models. Initially modelling activity largely followed the research of the Cowles Commission. The empirical macroanalysis used a large number of simultaneous equations, mostly linear and dynamic, which were statistically identified, while the role of economic theory was tantamount to the provision of pertinent regression variables for the equations (Klein, & Goldberger, 1955). The variables were often based on Keynesian IS-LM theories and the research focus was predominantly on their structural properties (feedback loops) and simultaneity, which was important for model estimation. The main disadvantage of the modelling approach propounded by the Cowles Commission was that while it was supposed to yield the dynamics resulting from the decisions of economic agents, the *ad hoc* equations failed to capture the mechanism of individual choice. Each variable was usually modeled with a separate equation and they were combined in modules, each separately worked out by a different group of experts. Then they were combined together and appended with an account of the interaction between the variables in different modules. These models were too compound, however, to provide a concise image of the mechanism, which propagated shock in the economy. They also failed to provide a reliable forecast of economic policy in a longer time perspective – a shortcoming less relevant for the Keynesian focus on short-term analysis and the question of how to effectively bring the economy back to equilibrium after the shock.

There are two major reasons for the failure of these models according to A. Spanos (1990, pp. 98-101), namely structural and statistical identification, due to the lack of appreciation of the intricate relation between theory and (statistical) measurement: "The theory determines not only the structural form and the relevant variables but the general form of the statistical model as well" (1990, p. 101). According to D. Hendry (1976) and D. Qin (1997, 2013) initially the *ex post* statistical identification was prioritized, with significantly less interest in the sources of the errors of the models. The problems with structural identification turned out, however, to be more fundamental as indicated by R. Lucas

[4] A general discussion of the concept of "paradigm" in macroeconomics is presented in (Galbács, 2015, pp. 2-6).

[5] For a comprehensive historical presentation see e.g. (Welfe, 2013; Vroey, 2016).

and C. Sims (Vroey, 2016). Lucas (1976) questioned the exogeneity of variables representing instruments of economic policy. While this structural identification did not explicitly take into account the expectations of economic agents, the 'structural' parameters of the model turned out to be a mixture of structural parameters and parameters related to the expectations of economic agents, which could not be assumed to be invariable for different regimes of economic policy as the estimation of those parameters changed with new economic policy. Therefore the model could not be used to predict the effects of economic policy change. Sims (1980) extended this criticism by pointing out that in the world of forward-oriented agents no variable can be consistently considered as exogenous. Thus endogeneity of economic policy leads to correlations between macroeconomic variables and variables-instruments of economic policy.

The crisis of the 1970's was decisive in changing attitudes among economists towards this kind of modelling. As succinctly expressed by (Pesaran & Smith, 1995, p. 66): "the models were seen as statistically *inadequate,* theoretically *inconsistent*, and practically *irrelevant*" as they "did not represent data", "did not represent theory" and "were ineffective for practical purposes of forecasting and policy analysis". Hence alternative approaches emerged, but at the early stage the most influential were the LSE method (Hendry, 1987) and SCVAR method (Structural Cointegrated Vector Autoregression) (Blanchard & Quah, 1989), which up to now are an important element of macroeconomic analysis. LSE modelling is based on the principle of reduction in order to cope with statistical identification. A model is understood as a representation of an unknown stochastic process, which generates the observable empirical evidence. The reduction consists in excluding variables on the basis of the statistical tests performed. The model is nonetheless required to be complete – if the statistical properties of the vector of the rest deviates from Gaussian white noise, then the model specification has failed.

Non-structural models of vector autoregression VAR are in fact a generalization of LSE models on vector time series. They express endogenous variables by means of lagged values. The VAR model of the order of $K \geq 1$ has the following form:

$$y_t = Ay_{t-1} + e_t,$$

where y_t is a vector of endogenous variables in time t, and A an autoregression matrix. The process e_t contains shocks, which influence the dynamics of endogenous variables and it gives ground to moving average representation. The assumption leading to structural interpretation of shocks (SVAR – structural VAR models) is that the matrix of covariance of shocks becomes unity. This approach has a more general form for non-stationary co-integrated processes (SCVAR – structural co-integrated VAR or SVECM – structural vector error correction). The advantage of these VAR models is that they attempt to solve both kinds of problems: structural and statistical identification. However the specification of VAR is not grounded theoretically and even for the structural versions the relationships between variables do not refer to any economic mechanism underlying the processes modelled. This results sometimes in non-intuitive responses to shocks and incoherent forecasts.

DSGE models[6] attempt to solve the problem of structural identification and therefore their specification is based on economic theory, usually by implementing the consumption Euler equation, the New Keynesian Phillips curve and the monetary policy rule, which are appended with the law of motion of economy (exogenous stochastic shocks). However, the reduced form of DSGE, which characterizes the dynamics of the endogenous variables of the model in long-term equilibrium, is a VAR model with some restrictions, which are motivated by the optimal decision rules of rational economic agents such as households, firms, government and the central bank.[7]

The structural version of DSGE model may be presented as follows:

$$\mathbb{E}_t \{f(y_{t+1}, y_t, y_{t-1}, \epsilon_t)\} = 0,$$

where ϵ_t is a vector of structural shocks and y_t a vector of all the other (endogenous) variables. The simple linearized model has the state space representation as a restricted VAR (Herbst & Schorfheide, 2016, p. 19):

$$s_t = \Phi_1(\theta)\, s_{t-1} + \Phi_\epsilon(\theta)\, \epsilon_t$$

with the coefficient matrices $\Phi_1(\theta)$ and $\Phi_\epsilon(\theta)$ as functions of the structural parameters of the DSGE model. The specification of the empirical model is complemented by measurement equations which enable the user to relate the parameters in s_t to a set of observables y_t. This is then fed into Bayes theorem together with the priors, which are usually further divided into three categories, in order to ground them in non-sample information (more details in Section 1).

[6] N. Kocherlakota succinctly explains the acronym 'DSGE': "Dynamic refers to the forward-looking behavior of households and firms. Stochastic refers to the inclusion of shocks. General refers to the inclusion of the entire economy. Finally, equilibrium refers to the inclusion of explicit constraints and objectives for the households and firms" (2010, pp. 9-10)

[7] The intuitive reading is discussed in (Tovar, 2009; Nachane, 2016)2016. For a more formal presentation see e.g. (Smets, & Wouters, 2003, 2007; Del Negro & Schorfheide, 2013; Christiano et al., 2010; Fernández-Villaverde, 2010; Lindé, Smets, & Wouters, 2016). For an argument that DSGE models can be improved to predict the 2008 financial crises see e.g. (Breuss, 2016).

1. Pragmatics in scientific representation

The methodological aspects of Romer's criticism frame the main line of his argument as he opens the paper with objections concerning the purported failure of DSGE models to meet scientific standards (2016, 1): "… the methods and conclusions of macroeconomics have deteriorated to the point that much of the work in this area no longer qualifies as scientific research". It is taken for granted that the standards are set up by physics:

> The evolution of macroeconomics mirrors developments in string theory from physics, which suggests that they are examples of a general failure mode of for fields of science that rely on mathematical theory in which facts can end up being subordinated to the theoretical preferences of revered leaders. The larger concern is that macroeconomic pseudoscience is undermining the norms of science throughout economics.

The envisioned consequence would have a wide-ranging impact on "all of the policy domains that economics touches".

My discussion in the reminder of this paper is organized by the first two main topics of Romer's methodological criticism concerning: i) objectivity, ii) empirical grounding and iii) diffusion of DSGE models (the latter only cursorily mentioned in the concluding Section 4). Although the charge regarding the lack of objectivity occupies less space in Romer's paper than the one on empirical grounding and is entangled within his remarks on the diffusion, it seems both more fundamental and devastating for DSGE models.

The lack of objectivity of DSGE models is implicitly addressed by Romer in the section of his paper titled "Questions about Economists, and Physicists" (2016, pp. 7-8). Having discussed problems with the empirical grounding of the models, he notes: "I find that it helps to separate the standard questions of macroeconomics, such as whether the Fed can influence increase the real fed funds rate, from meta-questions about how what [sic!] economists do when they try to answer the standard questions." It is the latter kind of question that guides the analogy between the apparent failure of the string theory in physics and "post-real" macroeconomics.[8] Romer puts the charge succinctly thus:

> The conditions for failure are present when a few talented researchers come to be respected for genuine contributions on the cutting edge of mathematical modeling. Admiration evolves into deference to these leaders. Deference leads to effort along the specific

[8] The analogy is only indicated, but not scrutinized in Romer's paper, so I do not enter into the details here, however, it should be remarked that it may not be a trivial task to conclusively demonstrate the "pseudo-scientific" character of string theory and perhaps even more difficult to undermine its role in advancing theoretical understanding in physics (Dawid, 2013).

> lines that the leaders recommend. Because guidance from authority can coordinate the efforts of other researchers, conformity to the facts is no longer needed as a coordinating device. As a result, if facts disconfirm the officially sanctioned theoretical vision, they are subordinated.

The "coordinating guidance of authority" is explained by reference to M. Bunge's distinction between "research" fields, such as mathematics, natural and engineering sciences, and "belief" fields, such as religion and political action. The latter are inherently characterized by the lack of "logical argument" and "fact" "that group members could use independently to resolve the question". If a "belief field" presents itself as a "research field" it qualifies as pseudoscience. The purported lack of objectivity would then consist of replacing the epistemic norms of "logical argument" and "fact", which are characteristic for "research fields", with "deference to authority".

Before we proceed to a more detailed discussion of this charge as to the lack of objectivity of DSGE models, a comment on the "coordinating guidance of authority" is needed. There are two different ways to interpret how "the coordinating guidance of authority" may relate to DSGE models as defined in Section 1. Romer predominantly focuses on authority as a factor *external* to the models themselves, determining – perhaps through social and psychological mechanisms – the kind of representation that becomes dominant among academics and decision-makers.[9] In that sense "deference to authority" compromises "conformity to the facts" and independent logical inference. On the other hand, it may well be that this authority derives from "genuine contributions on the cutting edge of mathematical modeling" as succinctly phrased by L. Christiano et al. (2010, p. 286): "These models have been shown to fit aggregate data well by conventional econometric measures. For example, they have been shown to do as well or better than simple atheoretical statistical models at forecasting outside the sample of data on which they were estimated. In part because of these successes, a consensus has formed (…)". So, the relationship between "authority" and the models could also be a factor, which is *internal* to the models themselves.[10] This internalist understanding is discussed below in more detail, but here it may be noted that it simply recognizes the salient fact that the perspective of the user of the model is itself constitutive for the model

[9] An indication that different adoption mechanisms operate for academia and decision-makers is presented in (Kocherlakota, 2010, p. 17).

[10] De Vroey & Pensieroso (2016) provide a comprehensive internalist account of the development of "mainstream" economics, where the driving factors are 'bifurcation decision nodes' (Leijonhufvud, 1994) related to theoretical choices. Other factors, fragmentation and certification, are conditional on the theoretical bifurcations. In particular the standardization of DSGE models is constituted by several methodological choices, such as "new mathematical tools allowing dynamic analysis", shift from "implicit" to "explicit microfoundations" and "a change from econometric testing to calibration first, and to Bayesian estimation later".

as scientific representation. While Romer's focus is obviously on the externalist interpretation, which is relevant to his view on the mechanism of diffusion of DSGE models, it seems rather untenable that he dismisses the epistemic merit of the internalist interpretation altogether.[11]

Turning now to the issue of objectivity, my contention throughout the paper is that recent advances in the theory of scientific representation (e. g. Giere, 2006; Wimsatt, 2007; van Fraassen, 2008; Ladyman, Bueno, Suárez, & van Fraassen, 2011; Mäki, 2013; Frigg & Nguyen, 2017) establish that the user's perspective is constitutive for *models as representations*. In my exposition I follow B. van Fraassen, but it may straightforwardly be extended to the others (Callebaut, 2012). On this "structural empiricist" view (Kawalec, 2016), roughly speaking, there are three components involved in scientific representation: phenomena (real world), appearances (data model) and theory (theoretical model). The latter has to be empirically grounded, which can be demonstrated by embedding the mathematical structure of data models within theoretical model. The data model is an outcome of the mathematical representation of the results of measurement. While the relationship of embeddedness between theoretical and data models may be well mathematically defined, it requires more elaboration to articulate when a data model *represents* the phenomena of interest. Put simply the problem is this: if abstract mathematical models, as noted by H. Weyl, can only represent up to isomorphism, how can we know that a given model represents the relevant concrete phenomena rather than a different set of phenomena, which may have an isomorphic structure? The answer is that it is "the indexical judgment" of the user in taking a particular mathematical model to represent the relevant phenomena. Hence scientific representation is an inherently user-related notion, which involves three – rather than just the first two – elements, namely: reality, its model and the user. Without the user's intention to represent a given domain of objects with a particular model, there is no scientific representation. So while scientific representation constitutively reflects the user's location and perspective, there is no scientific representation from "God's point of view". Scientific representation is based on the user's indexical judgment, reflecting her/his "here-and-now". Of course it is implicitly assumed by van Fraassen that what is meant by "user" is not just a particular individual, but rather a kind of "ideal-type" user in M. Weber's sense (Galbács, 2016, p. 21), implying that anyone with the necessary level of knowledge and skill, and having a particular location, will make the same indexical judgment. On the perspectival view, this intersubjective view of scientific representation may be the best we can make out of the ideal of objectivity. The modernist ideal

[11] This kind of radical approach to scientific representation characteristic of "the second wave" of science studies (Collins & Evans, 2002, 2009) has lost much of its bite under the severe criticism, especially by scientists themselves as epitomized by the "Sokal affair" and subsequent "waves" in science studies.

of God's view type of objectivity turns out to be untenable with regard to scientific representation, which is both perspectival and consensual.

I claim that a part of Lucas' critique (Lucas, 1976; Lucas, & Sargent, 1989) is tantamount to this point: the understanding of objectivity of scientific representation that the members of the Cowles Commission inherited from modern physics and embedded in the Klein-type of macroeconomic models disregarded the inherent perspective of the users. In Lucas' own conclusion:

> (…) given that the structure of an econometric model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any changes in policy will systematically alter the structure of econometric models (Lucas, 1976, p. 41).

It may therefore be concluded that the structural parameters, which distinguish DSGE models from the previous generation of models, but also from the contemporaneous VAR models (as discussed in more detail in Section 1), are identified by Lucas as a result of his taking into accout the user's perspective. If this account is correct then perspectivism of DSGE models cannot be taken as a *failure*, but rather as their constitutive element, which follows the pattern of scientific representation in general and is commonly considered as an advance in macroeconomic theorizing.[12] In that sense then "the internalist authority" legitimizing this perspective is not only consistent with epistemic norms of scientific representation, but is a clear advantage over the previous models, which as decisively pointed out by Lucas, confused God's view perspective with inter-subjectivity.

More specifically there are two ways in which DSGE models, according to Romer, fail to provide us with objective knowledge. One is the usage of the assumptions, which he calls "facts with unknown truth value (FWUTV)", and the other is the usage of Bayesian inference, in particular as to how the priors are set up (2016, p. 6). I will consider these in turn, arguing that they both follow the epistemic norms of empirical grounding applied in physics (as argued in Section 3), but here let me just shortly address one aspect of Bayesian priors.[13] The very use of priors makes the perspectivism of DSGE models explicit. It is most obvious in cases, where no evidence, such as nowcasts, earlier observations or micro-level data, is available to empirically ground the prior distribution of the model parameters $p(\theta)$. Then it is assumed that the available background knowledge X^0 – i.e. "nonsample

[12] For a detailed account of the advancements in mainstream economics see (de Vroey & Pensieroso, 2016).

[13] For an account of the frequentist estimation of DSGE see e.g. (Fernández-Villaverde, Ramírez, & Schorfheide, 2016).

(meaning other than Y) information" (Herbst, & Schorfheide, 2016, p. 24) – is independent of information in Y obtained from the investigated sample: $p(Y|X^0, \theta) \approx p(Y|\theta)$. On the standard epistemological understanding of the Bayesian principles of inference the user's background knowledge is not arbitrary, but reflects her/his *total evidence* (Williamson, 2002, pp. 200-201) preceding the research process, which would generate new data in Y to be included in the likelihood function. It is taken for granted that anyone with the same location as the user, having the same total evidence, will make the same judgment on priors. So if the requirement of total evidence is satisfied, then so is the requirement of inter-subjectivity.[14]

The perspectival character of DSGE models as scientific representation seems also implicitly recognized by Romer as exemplified in his discussion of a famous paper (Lucas, & Sargent, 1989). The latter quotes: "The problem of identifying a structural model from a collection of economic time series is one that must be solved by anyone who claims the ability to give quantitative economic advice" (Lucas & Sargent, 1989, p. 52 quoted by Romer 2016, p. 18). This statement concludes their discussion of macroeconomic modelling – not specifically DSGE – but structural models and in the preceding sentence they emphasize the constitutive user-relatedness of this kind of models:

> Without knowledge as to which structural parameters remain invariant as policy changes, and which change (and how), an econometric model is of no value in assessing alternative policies. It should be clear that this is true *regardless* of how well (3) and (4) fit historical data, or how well they perform in unconditional forecasting.

When combined with the above general characterization of perspectivism as constitutive for scientific representation, this claim indicates – in contrast to Romer's main objection – that the user's perspective has been recognized and acknowledged by economists at least since the famous Lucas critique of macro-modelling onwards. As indicated below there is, however, an important lesson to be learned by proponents of scientific perspectivism from the Lucas critique, namely that – at least in the case of social sciences – the user's perspective is *reflexive* and their "rational expectations" indicate an inherent two-directional relationship between the user's location and the model itself.[15]

[14] A similar point may also be elaborated for the classical frequentist statistical inference: "Of necessity, as it seemed to us (him and Neyman), we left in our mathematical model a gap for the exercise of a more intuitive process of personal judgement in such matters (…) as the choice of the most likely class of admissible hypotheses, the appropriate significance level, the magnitude of worthwhile effects and the balance of utilities (Pearson 1966, p. 277)" (quoted from (Howson & Urbach, 2005, p. 182)).

[15] In my earlier paper (Kawalec, 2016) I demonstrate it in more detail for the case of RBC models.

2. Empirical grounding

From the methodological perspective adopted here the bulk of Romer's paper focuses on the problems related to the empirical grounding of DSGE as an embodiment of macroeconomic theory.[16] Before discussing them in more detail I will outline the philosophical account of the criteria of empirical grounding and illustrate them with the historical example of advancements in atomic theory. Empirical grounding is the basic requirement for scientific research in general: "Parameters introduced into modeling must not be empirically superfluous – there must be, in some way, even if at some distance, a coordination with empirically differentiating phenomena" (van Fraassen, 2009, p. 10). In the problematic case of parameters which turn out to be empirically superfluous and cannot be thus "coordinated", the most straightforward solution would be to eliminate them. This, however, may not be feasible when the removal of the parameters concerned would affect the coherence or empirical strength of the theory: "The 'grounding' requirement turns into a salient problem only when elimination is not possible, while there are no theoretically specifiable conditions in which their values can be determined, relative to the theory, on the basis of measurement results" (van Fraassen, 2009, p. 10).

To grasp the solution to the "salient problem" of superfluous parameters which cannot be eliminated we first turn to observation concerning the role of theory in measurement. The objection raised by the Cartesians against Newtonian theory that it introduced "occult qualities" such as mass and force eventually led to the conclusion that "the measurement of those dynamic parameters on a body is an operation that *counts as such a measurement relative to Newtonian theory*" (van Fraassen, 2009, p. 9).[17] So it is necessary to take Newton's laws for granted in order to calculate the dynamic parameters from the directly measurable kinematic quantities, such as lengths and durations. Given the role of theory in measurement it then seems natural that: "The appropriate, and typical, response in that case is to start *enriching* the theory so that it becomes more informative, informative enough to allow the design of experiments in which this empirical determination of the values does become possible" (van Fraassen, 2009, p. 10).

To complete the account of empirical grounding we need a more detailed characterization of its criteria and the pitfalls of their application in the actual course of scientific research. Drawing upon the remarks of H. Weyl and their systematic exposition by C. Glymour, van Fraassen specifies three criteria of

[16] An overview of the main problems of DSGE is presented in (Blanchard, 2016), who nonetheless concludes: "I believe the DSGEs make the right basic strategic choices and the current flaws can be addressed" (p. 3).

[17] For an illustration of the measurement of mass by means of the Atwood machine see (van Fraassen, 2009, p. 9).

empirical grounding (2009, pp. 11-12): *determinability*, *theory-relativity* and *uniqueness*, where the two latter specify Weyl's criterion of *concordance*. The criterion of determinability is satisfied for any theoretically significant parameter if there are conditions under which measurement can determine its value. This determination presumes relationships, which are theoretically posited. There also must be a "unique coordination" of the parameter values, i.e. concordance between them,[18] when they are determined by different means.

It is instructive to go through the hundred years of the development of atomic theory and the kinetic theory of heat to trace the step by step theoretical and empirical advances aiming at the empirical grounding of those theories (van Fraassen, 2009, pp. 13-22). It was eventually accomplished by J. Perrin's experimental work on Brownian motion. The details of the whole story are beyond the scope of this paper, however, I need to highlight the critical moment. Starting in the 1820's J.-B. Dumas carried out research on the vapour densities of mercury and sulphur, which could have provided estimates for the atomic weights and could have improved the atomic theory. On the contrary, the measurements failed the criterion of concordance, leading to inconsistent determinations of the parameters. This, and similar problems, were addressed by Perrin by introducing into the theory more theoretically motivated constraints on the relationships between the parameters, which prior to that, resisted successful empirical grounding. Eventually it turned out that only one of the parameters still needed empirical grounding, which was accomplished by the study of the Brownian motion. Thus the first criterion of empirical grounding, namely determinability, was eventually met. By designing further experiments which linked to the then new Einstein theory, Perrin was also able to demonstrate concordance with the previous results.

The important lesson from this is that the empirical grounding of theoretical parameters is a matter of a dynamic interplay between the development of theoretical constraints on the parameters, on the one hand, and the design and performance of theoretically informed measurements on the other. Thus it definitely is not simply a matter of collecting data to confront an already finalized theory:

> This way we do not view it as a century-long search for independent evidence for the truth of a well-defined hypothesis about what nature is like, but in a quite different light. The enterprise of those scientists from Dalton to and including Perrin, aimed to develop the theory itself, and to enrich it so as to allow construction of models for special cases in its domain – all *so as to make empirical grounding possible* for its theoretical quantities (van Fraassen, 2009, p. 22).

[18] A more detailed specification of concordance is presented in (van Fraassen, 2009, p. 13).

Hence, when considering empirical grounding of theories, we have to bear in mind the inherently mutual relationship between measurements and theories: while evidence is always relative to theory, theory has to be informative enough to enable empirical testing. The task of the empirical grounding of a given theory may be considered complete if all its significant parameters are empirically grounded and the measurement results in different situations are in concord. This is what the criteria of empirical grounding discussed here imply.

With this outline of empirical grounding of scientific theories we can now turn to Romer's objections to the lack of empirical grounding of DSGE models (introduced in Section 2). They focus on two main issues: stochastic disturbances of the models and Bayesian estimation of the parameters. The shocks are called "imaginary" (2016, p. 4), because in contrast to the structural parameters of the DSGEs (Section 1), which reflect "actions that people take", they lack "microeconomic evidence" and "any sensible theoretical interpretation" of their meaning.[19] The identification problem arises, because "to get results, an econometrician has to feed something in other than observations on the variables in the system. I will refer to things that get fed in as facts with unknown truth value (FWUTV) (…)" (2016, p. 6).[20] This is predominantly achieved through Bayesian priors: "current practice in DSGE econometrics is to pin down estimates not simply by feeding in FWUTV by "calibrating" the values of some parameters, but increasingly through tight Bayesian priors for other parameters" (2016, p. 6). The problem proliferates as "the prior specified for one parameter will typically have a decisive influence on the output (a posterior distribution) for other parameters" (2016, p. 6).

Romer's criticism against "imaginary shocks" appears tantamount to the claim that they are empirically superfluous and that the way to proceed with them in macroeconomic theory is to dispense with them altogether, as clearly demonstrated by his analogy with phlogiston and similar concepts, recognized in science as empirically superfluous. The problem with empirical grounding or even the lack of clear meaning of the shocks notwithstanding, Romer, however, does not provide an argument to the effect that the proposed elimination will improve macroeconomic theorizing. Hence his proposed strategy may seem dubious insofar as stochastic disturbances, analyzed and computed based on the shocks to mimic actual aggregate fluctuations, form constitutive part of DSGE models, whose structural parameters yield an understanding of the mechanisms, which are fruitful for policy making (Tovar, 2009; Braun, 2017). *Prima facie* then we have good reasons to consider the suggested removal as

---

[19] N. Kocherlakota objects to the instrumental use of shocks as "convenient shortcuts to generate the requisite levels of volatility in endogenous variables" (2010, p. 16). V. Chari et al. claim that without a clear interpretation shocks impinge upon the structural character of the parameters in DSGE models (Chari, Kehoe, & McGrattan, 2009, p. 243).

[20] For a systematic exposition of the notion of identification see e.g. (Rothenberg, 1971; Reicher, 2016, p. 414).

a step diminishing the theoretical coherence of macroeconomic theory as embedded in the DSGEs and thus violating the criteria of empirical grounding.

The identification problem is tackled by Romer in recognition of two different aspects, which correspond to the two strategies observed in the historical case of the empirical grounding of atomic theory. The first one is related to bringing in new empirical evidence to determine the parameter values. The second concerns the elaboration of theoretical constraints on the parameters to increase their determinability. Romer is critical of the latter, especially the proposed introduction of rational expectations in order to set "cross-equation restrictions" (Lucas & Sargent, 1989). He objects to this solution as it will increase the number of parameters, but primarily because "to make causal inferences, the econometrician has to feed in some facts that are known to be true and math cannot establish the truth value of a fact. Never has. Never will. In practice, what math does is let macroeconomists hide the FWUTV's that they feed into their estimation procedure" (2016, p. 14).

This skepticism seems much like Dumas's rejection of the atomic theory in the 1830's in the light of the negative outcomes of his experiments. However, what turned out to be productive in that case was in fact more theoretical effort, which then stimulated and enabled the design of empirical measurements that finally resolved the problematic issue. Against this background Romer's conclusion may appear a *non sequitur*: given the historical record of attempts to empirically ground theories in science it seems that *more theoretical effort* is needed to enable collection of empirical evidence to decide the case, either way.[21]

Romer's preferred solution would be to provide more empirical evidence: "Do experiments. For macroeconomists, this means looking for natural experiments" (2016, p. 13). As outlined earlier, however, the relevant evidence counts as such *only relative to theory*. Let me repeat the moral: empirical grounding can only be achieved by mutual adjustments of theory and measurement. It is a distorted view of the empirical assessment of theory that the latter would be ready-made and as such would merely await collection of data (Sims, 1996, p. 106). For what counts as data depends on the elaboration of a sufficiently informative theory, which enables the design of pertinent measurements.

Admittedly the problems of DSGE models with identification are very difficult.[22] One source of the problems is the inability to distinguish between endogenous and exogenous sources of persistence in DSGEs, for instance in case where it is due to price adjustment costs vs the case where the exogenous shocks are highly correlated. Another source of problems arises due to

[21] With a few exceptions the historical cases of recognition of failure – starting from Ptolemaic epicycles through phlogiston and Klein-type models – resulted from effort along the two lines mentioned by Romer himself: empirical evidence and logical inference, which eventually established better alternatives.

[22] For an instructive debate between some prominent economists see (Solow, 2010).

the rational expectations. For instance, if the observed volatility of inflation is very low, then it is difficult to precisely determine the coefficient of policy rule (Herbst, & Schorfheide, 2016, p. 78). Moreover, the generalization of the law of motion, which is needed in serious applications of DSGEs, such as central bank forecasting and which amplifies the problems with identification: "The more flexible and densely parameterized the law of motion of the exogenous shocks, the more difficult it becomes to identify the shock parameters and the parameters associated with the endogenous propagation mechanism jointly" (Herbst & Schorfheide, 2016, p. 132).

The main source of the problems with identification may be explained thus: "this problem occurs for the simple reason that a model with an unrestricted VAR shock process can approximate an unrestricted VAR in the observables arbitrarily well for any valid set of deep parameter values. Therefore, the concentrated likelihood function is completely flat with respect to those parameters, and those parameters hence are unidentified" (Reicher, 2016, p. 413).[23]

Do these shortcomings suffice to conclusively reject DSGE models as unscientific? First, the problems of identification pertain to a wide class of models not only in economics, but throughout the social sciences. In that sense the scientific status of DSGE models is then on a par with a major part of social science research. Apparently though Romer's argument is not intended for such a broad scope of models. Second, these problems with identification stimulate a broad range of theoretical and empirical research in macroeconomics and perhaps – as Dumas' case may suggest – this will eventually reach a climax, where an informed decision could be made concerning the fate of DSGE models.

The other problem with identification that Romer discusses is the use of Bayesian priors.[24] One objection is that priors determine the posterior distribution – using different priors may change the posterior distribution regardless of the likelihood function and the data: "By changing the priors I feed in for the supply curve, I can change the posteriors I get out for the elasticity of demand until I get one I like" (2016, p. 15). Thus the determinability criterion of empirical grounding, discussed earlier, will be compromised. Romer also accuses Bayesian priors of failing to meet concordance. For instance with regard to the SW model (Smets & Wouters, 2003) he refers to papers discussing how the identification and estimation of the structural estimates depends on the priors.

The use of Bayesian priors is generally explained thus: "They play an important role in the estimation of DSGE models because they allow researchers to incorporate information not contained in the estimation sample into the em-

[23] A formal proof that without constraints on shocks the parameters in θ are not identifiable is presented in (Reicher, 2016, pp. 417-418).

[24] The application of the principles of Bayesian inference to DSGE models is discussed in detail in (An & Schorfheide, 2007; Herbst & Schorfheide, 2016).

pirical analysis" (Herbst & Schorfheide, 2016, p. 22), which turned out to be more attractive than the traditional calibration (Kawalec, 2016). In response to the problem of the lack of empirical grounding of the priors it is also acknowledged that: "While priors could in principle be formed by pure introspection, in reality most priors (as well as most model specifications) are based on some empirical observations." (An & Schorfheide, 2007, p. 24). One way to enhance the empirical determinability of the values of the priors is to group them into the three sets of the structural parameters in θ: parameters affecting the steady state of the DSGE model (e.g. the interest rate, inflation rate, growth rate), parameters characterizing the law of motion of the exogenous shocks and parameters that control the endogenous propagation of mechanisms. Parameters in the first category are based on pre-sample averages, for instance, if the estimation sample begins with Q1 of 1983, then the priors may be based on data from the 1970's. As shock processes are unobserved, the priors in the second category are elicited indirectly: "beliefs – possibly informed by pre-sample observations – about the volatility, autocorrelations, and cross-correlation of output growth, inflation, and interest rates, could be mapped into beliefs about the persistence and volatility of the exogenous shocks" (Herbst, & Schorfheide, 2016, p. 25). Finally, the third group of parameters uses micro-level data, for example for price adjustment or labour supply elasticity.[25]

One of the advantages of using Bayesian inference are convergence results. An early result was proven in (Halpern, 1974) for a linear regression model – when selected at random from the class of such models that generate the data, its posterior probability will converge to 1 as the number of observations goes to infinity. This result then was applied (Fernández-Villaverde & Rubio-Ramírez, 2004) to DSGE models with misspecification. In addition Bayesian model averaging was developed to use a mixture of distributions by averaging across models with their posterior probabilities as weights (Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, & Santaeulàlia-Llopis, 2012; Del Negro & Schorfheide, 2013, pp. 131-132).

Of course a number of problems remain. For instance, "the selection of certain priors may produce outcomes that look good from a theoretical perspective, even if the data is mute about the parameter value" (Tovar, 2009, p. 15). In other words, "arbitrarily chosen priors may hide severe identification problems" (Canova & Sala, 2009, p. 432). This is for instance true of the standard SW model, where "the estimates of these key parameters are very sensitive to the prior distribution" (Herbst & Schorfheide, 2016, p. 142). Another problem for Bayesian estimation is that the replicability of posterior estimates may be

[25] Usually, the determination of prior distributions is an iterative procedure to ensure that the distributions for individual parameters do not lead to implausible consequences for the joint prior distribution; for details see (Herbst & Schorfheide, 2016, pp. 25-26).

difficult due to reliance on computationally intensive simulation methods, such as the Metropolis-Hasting algorithm (Tovar, 2009, p. 16).

It has been demonstrated for ideal agents that the only coherent epistemic policy of updating beliefs is Bayesian updating (Diaconis & Zabell, 1982; van Fraassen, 1999). More recently, it has also been demonstrated that a condition – called "Tracking" – which generalizes Bayesian updating, holds for a wider class of non-Bayesian updating rules (van Fraassen & Halpern, 2016). The motivation behind the tracking criterion is consistent with the concept of empirical grounding discussed earlier in this section: "opinion represented in terms of a probability assignment should at least possibly track the relevant statistics, and updating the probability assignment on new input should preserve that possibility" (van Fraassen & Halpern, 2016, p. 4). Of course, it may turn out that there are good reasons to depart from the Bayesian updating altogether, or even the generalized tracking criterion, in the case of macroeconomic theorizing, as for instance in a pertinent account of "tipping points" in the economy. But given our current evidence there are no grounds to consider Bayesian inference as epistemically inferior to alternative approaches to the extent that it would qualify as pseudo-scientific.

Conclusions

I conclude the paper by some observations related to the last part of Romer's criticism, namely the purported "imperialism" of DSGE models as enforced by "deference to authority" in the externalist sense discussed in Section 2. A satisfactory response to this kind of argument, and specifically an evaluation of the extent to which a parallel development of alternative models could have been impeded, would require extensive empirical research into the diffusion and adoption of DSGE models by both academia and banks.[26]

Admittedly it was the advancements from RBC to DSGE, and in particular the SW models elaborated for the European Central Bank, that have significantly contributed a more widespread adoption of DSGE models in analysis and forecasting (Tovar, 2009; Kocherlakota, 2010, p. 17). Of course DSGE models may be outperformed in different aspects, such as short-term forecasting. Nonetheless their overall performance seems much better than other alternatives: "While a successful decathlete may not be the fastest runner or the best discus thrower, he certainly is a well-rounded athlete" (Del Negro & Schorfheide, 2013, p. 61).

[26] Cf. e.g. (Sanbonmatsu, Posavac, Behrends, Moore, & Uchino, 2015). Also remarks by prominent proponents of DSGE models, such as (Fernández-Villaverde, 2010, p. 5), are less then fully substantiated by the evidence.

References

An, S., & Schorfheide, F. (2007). Bayesian analysis of DSGE models. *Econometric Reviews*, *26*(2-4), 113-172.

Blanchard, O. (2016). *Do DSGE models have a future?* (No. PB 16-11). Peterson Institute for International Economics.

Blanchard, O. J., & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, *79*(4), 655-673.

Braun, B. (2017). Central bank planning? Unconventional monetary policy and the price of bending the yield curve. In J. Beckert, & R. Bronk (Eds.), *Uncertain futures: imaginaries, narratives, and calculation in the economy* (p. forthcoming). Cambridge: Cambridge University Press.

Breuss, F. (2016). *Would DSGE models have predicted the great recession in Austria?* (No. 530/2016). WIFO.

Callebaut, W. (2012). Scientific perspectivism: A philosopher of science's response to the challenge of big data biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 69-80.

Canova, F., & Sala, L. (2009). Back to square one: Identification issues in DSGE models. *Journal of Monetary Economics*, *56*(4), 431-449.

Chari, V. V., Kehoe, P. J., & McGrattan, E. R. (2009). New Keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, *1*(1), 242-266.

Christiano, L. J., Trabandt, M., & Walentin, K. (2010). DSGE models for monetary policy analysis. In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics* (Vol. 3, pp. 285-367). Amsterdam: Elsevier.

Collins, H. M., & Evans, R. (2002). The third wave of science studies: Studies of expertise and experience. *Social Studies of Science*, *32*(2), 235-296.

Collins, H. M., & Evans, R. (2009). *Rethinking expertise*. Chicago: Univ. of Chicago Press.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge: Cambridge University Press.

de Vroey, M. (2016). *A history of macroeconomics from Keynes to Lucas and beyond*. New York: Cambridge University Press.

de Vroey, M., & Malgrange, P. (2010). From The Keynesian revolution to the Klein-Goldberger model: Klein and the dynamization of Keynesian theory. *University of Louvain, Department of Economics, Discussion Paper*, (2010-19).

de Vroey, M., & Pensieroso, L. (2016). *The rise of a mainstream in economics* (Discussion Paper No. 2016-26). Louvain: University of Louvain, Department of Economics.

Del Negro, M., & Schorfheide, F. (2013). DSGE model-based forecasting. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, pp. 57-140). Amsterdam: Elsevier.

Diaconis, P., & Zabell, S. L. (1982). Updating subjective probability. *Journal of the American Statistical Association*, *77*(380), 822-830.

Fernández-Villaverde, J. (2010). The econometrics of DSGE models. *SERIEs*, *1*(1-2), 3-49.

Fernández-Villaverde, J., & Francisco Rubio-Ramírez, J. (2004). Comparing dynamic equilibrium models to data: a Bayesian approach. *Journal of Econometrics*, *123*(1), 153-187.

Fernández-Villaverde, J., Ramírez, J. F. R., & Schorfheide, F. (2016). *Solution and estimation methods for DSGE models*. Cambridge, MA: National Bureau of Economic Research.

Frigg, R., & Nguyen, J. (2017). Scientific representation is representation-as. In H.-K. Chao, & J. Reiss (Eds.), *Philosophy of science in practice* (pp. 149-179). Cham: Springer.

Galbács, P. (2015). *The theory of new classical macroeconomics: a positive critique*. Cham: Springer.

Galbács, P. (2016). Beyond the realism of mainstream economic theory. Phenomenology in economics. *Economics and Business Review*, *2*(4), 3-24.

Giere, R. N. (2006). *Scientific perspectivism*. Chicago: University of Chicago Press.

Grabek, G., Kłos, B., & Koloch, G. (2010). *SOE PL 2009–Model DSGE małej otwartej gospodarki estymowany na polskich danych. Specyfikacja, oceny parametrów, zastosowania* (Materiały i Studia No. 251). Warszawa: Narodowy Bank Polski. Retrieved from http://pki.nbp.pl/publikacje/materialy_i_studia/ms251.pdf

Halpern, E. F. (1974). Posterior consistency for coefficient estimation and model selection in the general linear hypothesis. *The Annals of Statistics*, *2*(4), 703-712.

Hendry, D. F. (1976). The structure of simultaneous equations estimators. *Journal of Econometrics*, *4*(1), 51-88.

Hendry, D. F. (1987). Econometric methodology: a personal perspective. In T. F. Bewley (Ed.), *Advances in econometrics* (pp. 29-48). Cambridge: Cambridge University Press.

Herbst, E. P., & Schorfheide, F. (2016). *Bayesian estimation of DSGE models*. Princeton: Princeton University Press.

Howson, C., & Urbach, P. (2005). *Scientific reasoning: the Bayesian approach* (3rd ed). Chicago: Open Court.

Kawalec, P. (2016). Interaction and structural representation in calibration of economic models. *Studia Metodologiczne*, *36*(4), 131-145.

Klein, L. R., & Goldberger, A. S. (1955). *An econometric model for the United States, 1929-1952*. Amsterdam: North Holland.

Kocherlakota, N. (2010). Modern macroeconomic models as tools for economic policy. *The Region*, (May), 5-21.

Ladyman, J., Bueno, O., Suárez, M., & van Fraassen, B. C. (2011). Scientific representation: A long journey from pragmatics to pragmatics: Bas C. van Fraassen: Scientific representation: Paradoxes of perspective. Oxford: Clarendon Press, 2008. *Metascience*, *20*(3), 417-442.

Leijonhufvud, A. (1994). Hicks, Keynes, and Marshall. In H. Hagemann & O. F. Hamouda (Eds.), *The legacy of Hicks: his contribution to economic analysis* (pp. 147-162). London: Routledge.

Lindé, J., Smets, F., & Wouters, R. (2016). Challenges for central banks' macro models. In C. R. Taylor & H. Uhlig (Eds.), *Handbook of macroeconomics* (Vol. 2, pp. 2185-2262). Amsterdam: Elsevier.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, *1*, 19-46.

Lucas, R. E., & Sargent, T. (1989). After Keynesian macroeconomics. In F. E. Morris (Ed.), *After the Phillips curve: persistence of high inflation and high unemployment.*

Proceedings of a Conference Held at Edgartown, Massachusetts June 1978 (pp. 49-72). Boston: The Federal Reserve Bank of Boston.

Mäki, U. (2013). Contested modeling: the case of economics. In U. Gähde, S. Hartmann, & J. H. Wolf (Eds.), *Models, simulations, and the reduction of complexity* (pp. 87-106). Berlin, Boston: de Gruyter.

Nachane, D. (2016). Dynamic stochastic general equilibrium (DSGE) modelling in practice: identification, estimation and evaluation. *Macroeconomics and Finance in Emerging Market Economies*, *10*(2), 1-28.

Pearson, E. S. (1966). Some thoughts on statistical inference. In *The selected papers of E.S. Pearson* (pp. 276-183). Cambridge: Cambridge University Press.

Pesaran, M. H., & Smith, R. (1995). The role of theory in econometrics. *Journal of Econometrics*, *67*(1), 61-79.

Qin, D. (1997). *The formation of econometrics: a historical perspective* (1. paperback ed). Oxford: Clarendon Press.

Qin, D. (2013). *A history of econometrics: the reformation from the 1970s*. Oxford: Oxford University Press.

Reicher, C. (2016). A note on the identification of dynamic economic models with generalized shock processes. *Oxford Bulletin of Economics and Statistics*, *78*(3), 412-423.

Ríos-Rull, J.-V., Schorfheide, F., Fuentes-Albero, C., Kryshko, M., & Santaeulàlia-Llopis, R. (2012). Methods versus substance: Measuring the effects of technology shocks. *Journal of Monetary Economics*, *59*(8), 826-846.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, *39*(3), 577-591.

Romer, P. (2016). *The trouble with macroeconomics*. Commons.

Sanbonmatsu, D. M., Posavac, S. S., Behrends, A. A., Moore, S. M., & Uchino, B. N. (2015). Why a confirmation strategy dominates psychological science. *PLOS ONE*, *10*(9), e0138197. https://doi.org/10.1371/journal.pone.0138197

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1-48.

Sims, C. A. (1996). Macroeconomics and methodology. *The Journal of Economic Perspectives*, *10*(1), 105-120.

Smets, F., & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European Economic Association*, *1*(5), 1123-1175.

Smets, F., & Wouters, R. (2007). Shocks and frictions in US business cycles: A bayesian DSGE approach. *The American Economic Review*, *97*(3), 586-606.

Solow, R. (2010). Building science for a real world. Testimony presented at a hearing before the Subcommittee on Investigations and Oversight, Committee on Science and Technology (US House of Representatives). Retrieved from https://www.gpo.gov/fdsys/pkg/CHRG-111hhrg57604/pdf/CHRG-111hhrg57604.pdf.

Spanos, A. (1990). The simultaneous-equations model revisited. *Journal of Econometrics*, *44*(1-2), 87-105.

Tovar, C. E. (2009). DSGE models and central banks. *Economics: The Open-Access, Open-Assessment E-Journal*, *3*(2009-16), 1-32. https://doi.org/10.5018/economics-ejournal.ja.2009-16.

van Fraassen, B. C. (1999). Conditionalization, a new argument for. *Topoi*, *18*(2), 93-96.

van Fraassen, B. C. (2008). *Scientific representation: paradoxes of perspective*. New York: Oxford University Press.
van Fraassen, B. C. (2009). The perils of Perrin, in the hands of philosophers. *Philosophical Studies*, *143*(1), 5-24.
van Fraassen, B. C., & Halpern, J. Y. (2016). Updating probability: tracking statistics as criterion. *The British Journal for the Philosophy of Science*, axv027. https://doi.org/10.1093/bjps/axv027.
Welfe, W. (2013). *Macroeconometric models*. Berlin: Springer.
Williamson, S. (2017, January). The trouble with Paul Romer. Retrieved January 25, 2017, from http://newmonetarism.blogspot.com/2017/01/the-trouble-with-paul-romer.html
Williamson, T. (2002). *Knowledge and its limits*. Oxford: Oxford University Press.
Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: piecewise approximations to reality*. Cambridge, MA: Harvard University Press.