Do regional economists answer the right questions?

On the discrepancy between the questions regional economists solve and the questions policy makers actually ask

Thomas de Graaff[1]*

Abstract
In this position paper I argue two things: namely, (*i*) regional (or spatial) economists do not always answer the type of questions policy makers are concerned about, and consequently (*ii*) they are very restrictive in the tool set they apply. First, policy makers—whether national, regional or local—are oftentimes concerned about holistic approaches and future predictions. Exemplary questions are "What works best for my country/region/city" and "If we change this what will happen to the country/region/city as a whole?". Regional economists—actually, most economists—usually isolate phenomena in order to, at best, explain the impact of a single determinant. Indeed, most regional economists feel very uncomfortable when asked to predict or give the best set of determinants for a certain phenomenon. This has its consequences for the tool set the regional economists applies. Usually a parametric regression type of framework is applied isolating the determinant under consideration and controlling as much as possible for observables and unobservables, ideally in a pseudo-experimental framework. A direct consequence of this approach is that emphasis is very much on explaining the impact of a determinant and not on predicting phenomena. For many applications that is definitely the right approach. However, as this paper ultimately argues, it is very much as well a selective approach that does not do well to deliver on some of the questions policy makers ask regional economists.

Keywords
Regional science — spatial heterogeneity — conditional robustness — predicting — data science

[1] *Department of Spatial Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*
***Corresponding author**: ✉ t.de.graaff@vu.n; thomasdegraaff.nl

Contents

Introduction

> The sexiest job in the next 10 years will be statisticians.
>
> Hal Varian, 2009

The above quote from Hal Varian is in one aspect wrong; nowadays, we do not call them statisticians but data scientists instead. Nevertheless, in the last two decades companies such as Google, Ebay, Whatsapp, Facebook, Booking.com and Airbnb, have not only witnessed enormous growth but also changed the socio-economic landscape. Indeed, with the increasing abundance of (spatial) data and computer capacity, the ability to gather, process, and visualize data has become highly important and therefore highly in demand. And the models and tools these data scientists within these companies use are very much *data driven* with often remarkable results.

Therefore it is striking that in economics in general, and in regional economics in specific, most of the tools employed are very much *theory driven* instead of data driven. My (conservative) estimate would be that 90% of all empirical work in regional economics revolves around postulating a (linear) model and testing whether (a) key determinant(s) are significantly different from a hypothesized value—usually zero.[1] That is, *within* the context of the model assumed.

At best, this approach can be seen in a causal inference framework. If a determinant (policy) x changes, does it cause

[1] In a seminal contribution, Breiman, 2001 states that deep into the 90s 95% of the statisticians actually employed the theory driven paradigm and only 5% a data driven paradigm. With the advent of the availability of internet connectivity, large (online) data sources, and faster computers the statistical realm changed dramatically. However, this has not permeated yet in the social sciences.

then a change in the output (welfare) y?.[2] This approach provides a rigid and useful approach to regional policy evaluation.

1. Background

Regions are conceptually different than cities, as they contain urban, suburban and rural areas simultaneously. Whilst smaller regions can still be seen as the total influence radius of metropolitan areas—such as measured by the concept of local labor markets and the NUTS-2 regions in Europe—, larger regions can typically contain a multiple of cities in combinations with their various hinterlands –such as the Dutch Randstad, the Belgian Flemish Diamond, and the German Ruhr areas.

In recent years, the urban economics literature witnessed large growth; not only noticed by a wider acceptance in mainstream economics, but as well by larger scientific rigor and increased robustness of empirical findings. Remarkably, the empirical regional economics (or regional science in general for that matter) literature lagged behind, although many concepts and challenges in both disciplines are conceptually similar and are derived from similar theoretical backgrounds.

Similar to urban economics, there is not yet a clear (consensus in) understanding in which policy instruments are actually (cost-)effective in promoting regional growth.

To do so, I first review the previous literature in section 1. This section focuses mainly on regional economics as it has a larger emphasis on *causal* effects. To a lesser extent we deal with the (economic geography) literature. Based on this literature review Section 2 deals with the research gaps that can be identified.

- housing & population;
- amenities;
- connectivity & accessibility
- networks;
- social & humn capital

2. Research gaps

3. Research Agenda

3.1 Regional heterogeneity

(Thissen et al., 2016; de Graaff et al., 2012b,a)

3.2 Conditional robustness

In regional science in general and in regional economics in specific, remarkably little attention has been given to reproducibility and robustness of results (with some exceptions as, amongst some others, by Rey, 2014; Arribas-Bel and de Graaff, 2015; Arribas-Bel et al., Forthcoming).

[2] Most of this research actually intends to mimic a *difference-in-difference* approach

3.3 Regional sorting models

As in Bayer et al. (2004) and Bayer and Timmins (2007) and recently by Zhiling et al. (2016) and Bernasco et al. (Forthcoming).

4. In conclusion

References

Arribas-Bel, D. and T. de Graaff (2015). "WooW-II: Workshop on open workflows". In: *REGION* 2.2, pp. 1–2.

Arribas-Bel, D., T. de Graaff, and S. Rey (Forthcoming). "Looking at John Snow's cholera map from the XXIst Century: a practical primer on reproducibility and Open Science". In: *Regional Research Frontiers: The Next 50 Years*. Ed. by R. Jackson and P. Schaeffer. Berlin: Springer.

Bayer, P. and C. Timmins (2007). "Estimating Equilibrium Models Of Sorting Across Locations". In: *The Economic Journal* 117.518, pp. 353–374.

Bayer, P., R. McMillan, and K. Rueben (2004). "An Equilibrium Model of Sorting in an Urban Housing Market". NBER working paper: No. w10865.

Bernasco, W., T. de Graaff, J. Rouwendal, and W. Steenbeek (Forthcoming). "Social Interactions and Crime Revisited: An Investigation Using Individual Offender Rates". In: *Review of Economics and Statistics*.

Breiman, L. (2001). "Statistical Modeling: The Two Cultures". In: *Statistical Science* 16.3, pp. 199–231.

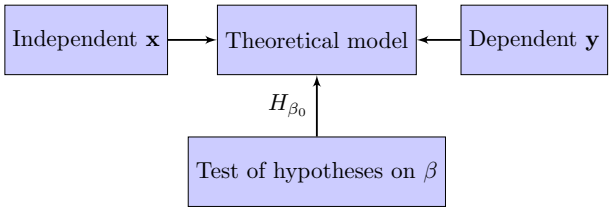
De Graaff, T., F. G. van Oort, and R. J. Florax (2012a). "Regional Population-Employment Dynamics Across Different Sectors of the Economy". In: *Journal of Regional Science* 52.1, pp. 60–84.

— (2012b). "Sectoral heterogeneity, accessibility and population-employment dynamics in Dutch cities". In: *Journal of Transport Geography* 25, pp. 115–127.

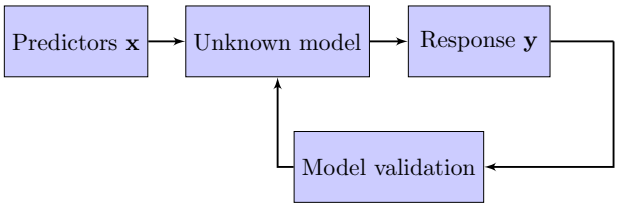
Rey, S. J. (2014). "Open regional science". In: *Annals of Regional Science* 52.3, pp. 825–837.

Thissen, M., T. de Graaff, and F. G. van Oort (2016). "Competitive network positions in trade and structural economic growth: A geographically weighted regression analysis for European regions". In: *Papers in Regional Science* 95.1, pp. 159–180.

Zhiling, W., T. de Graaff, and P. Nijkamp (2016). "Cultural Diversity and Cultural Distance as Choice Determinants of Migration Destination". In: *Spatial Economic Analysis* 11.2, pp. 176–200.



(a) Model approach proces



(b) Data approach proces

Figure 1. Two cultures of statistical/econometric modeling (inspired by Breiman, 2001)