

A Service of



Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Siliverstovs, Boriss; Kholodilin, Konstantin

Conference Paper Measuring Regional Inequality by Internet Car Price Advertisements: Evidence for Germany

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Siliverstovs, Boriss; Kholodilin, Konstantin (2012) : Measuring Regional Inequality by Internet Car Price Advertisements: Evidence for Germany, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at: https://hdl.handle.net/10419/120721

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

Measuring Regional Inequality by Internet Car Price Advertisements: Evidence for Germany

Konstantin A. Kholodilin DIW Berlin Mohrenstraße 58 10117 Berlin Germany kkholodilin@diw.de Boriss Siliverstovs ETH Zurich KOF Swiss Economic Institute Weinbergstraße 35 8092 Zurich, Switzerland boriss.siliverstovs@kof.ethz.ch

February 13, 2012

Abstract

We suggest to use car sale prices from internet advertisements for measuring economic inequality between and within German regions. Our estimates of regional income levels and Gini indices based on advertisements are highly positively correlated with the official figures.

Keywords: Car prices; economic inequality; German NUTS1 and NUTS2 regions; Gini index; Internet JEL code: C21; O47; R11.

1 Introduction

Economic inequality of households is an important characteristic of the welfare of a country or region. Societies experiencing intolerably high levels of inequality might be subject to more criminality, increased drug and alcohol consumption, as well as political instability. Governments usually try to avoid such scenarios at all cost. Government policies to combat inter-regional inequality include fiscal federalism and structural policy. However, governments need certain indicators to measure their success in this respect. These indicators are typically based on survey data. Data are collected from a limited number of representative households, which are asked to fill out questionnaires which consist of various questions concerning expenditure and incomes of the household. While such a practice of data collection is widespread, there are a number of problematic issues: 1) Only a limited number of households are selected (invited) to participate, which makes the data not representative at the regional level. For example, in Germany household surveys are mainly conducted by the Socio-Economic Panel (SOEP), which typically collects the responses of approximately 11,000 households. Provided that there are 16 NUTS1 and 39 NUTS2 regions in Germany, that would imply that on average there are less than 700 and 200 observations per region, respectively. 2) Participation is voluntary and verification of supplied information is costly and may not always be possible. 3) Voluntary surveys suffer from the so-called "middle-class bias" (Becker and Hauser, 2003) when households with very high and very low income levels typically are not sufficiently well represented.

In this paper we suggest to use information contained in car sale prices from advertisements placed on the web for measuring economic inequality in Germany at the national level and within regions. Our paper is motivated by the following considerations. First, the Internet offers an enormous amount of information. Compared with traditional information collection methods, an important advantage of using Internet data is that these can be retrieved much faster, practically providing real-time monitoring of ongoing processes in society. Thus, the Internet can be viewed as an important source of complementary information to traditional sources. In a pioneering study that appeared online in November 2008, Ginsberg et al. (2009) shows how one can use disaggregated Google searches filed by millions of users each day in order to study the intensity of influenza activity in the US. Later several studies have advocated the use of web queries for forecasting unemployment (e.g., Askitas and Zimmermann, 2009; Choi and Varian, 2009; D'Amuri and Marcucci, 2009) and private consumption (Kholodilin et al., 2010).

Second, our choice of using cars for capturing economic inequality between regions is not purely incidental. Cars is a specific durable good used not only for transportation but also for signalling social status of its owner and his economic well-being. In this respect, cars are different from other durable goods like fridges or washing machines that are primarily used for what they are built for.

Third, in many respects Germany is an optimal example for our study. It is geographically and economically a diverse country, where government policy on lessening regional economic inequality have been on the top of the political agenda since re-unification. Germans, as a car-building nation, love their cars and the well-maintained network of German highways makes it pleasant to drive them. Moreover, a car is an affordable good in Germany. Indeed, in a country with 39.8 million households there are about of 50.2 million registered cars (Federal Motor Transport Authority, (Kraftfahrt Bundesamt, (KBA))) and 81.2% of German households possess one or more cars according to the 2008 SOEP survey. In addition, it is also important for our approach that the Internet is widely available to German households. Accoring to Eurostat, in 2009 about 71% of individuals had regular access to the Internet in Germany.

2 Data

The data were downloaded from two specialised websites: http://www.mobile.de and http://www.autoscout24.de in December 2011. The corresponding number of unique car price observations downloaded from each website is 914,105 and 802,047, respectively. The following information were recorded: make, model, ZIP code, mileage, engine volume in liters and cubic centimeters, type of transmission (manual, automatic, etc.), the year of first registration, and offer price. The ZIP code information were used to find the geographical coordinates (latitude and longitude) of each car's seller. Then the price data were assigned to the respective NUTS1 and NUTS2 regions, given information about their borders. A shapefile containing geographical information on the regional borders was taken from Eurostat.

3 Results

Table 1 presents the results. The first row of the table reports the correlation coefficient for the number of advertisements recorded in the respective NUTS1 or NUTS2 regions and the number of inhabitants. At both levels of disaggregation, the correlation is close to unity, implying that the distribution of advertisements across German regions is proportional to the population in these regions, and hence our sample is representative.

The middle panel of Table 1 reports the estimated correlation coefficient between different measures of income and average car prices recorded for the respective NUTS1 or NUTS2 regions. For mobile.de, the highest correlations of 0.861 and 0.858 are reported for the balance of primary income and national income (wage and property income) both measured in per capita terms at the NUTS1 level; see Figures 1 and 2.¹ For autoscout24.de the magnitude of the correlation coefficients exceeds 0.8 for all measures of income at the NUTS1 level. The maximum correlation of 0.901 is reported for disposable income. Using data from mobile.de, the correlation coefficient between the average car price and the balance of primary income computed for the NUTS2 regional classification is 0.795; see Figures 3 and 4.² For autoscout24.de the corresponding correlation coefficient is slightly lower, 0.694. The observed high correlation between different income measures and average car prices indicates that lower income levels in the Eastern German federal states seem to be well reflected in lower average car prices. This supports our assumption that car prices observed in a given region could serve as a good proxy for regional income levels, i.e. households living in a poor neighbourhood tend to demand on average cheaper cars compared to households that live in better-off

 $^{^{1}}$ In order to save space we report only graphical results for car prices from mobile.de. We make the figures based on autoscout24.de available upon request.

 $^{^{2}}$ Abbreviations of the NUTS2 regions is according to Eurostat (2011).

regions.

We compare the values of the Gini index computed based on car prices with the official values based on net equivalence income (Statistisches Bundesamt, 2011) available for the NUTS1 regions. The estimated correlation coefficient is 0.809 for price data collected from mobile.de and 0.768 for autoscout24.de, reflecting close correspondence between these two measures; see Figures 5 and 6.

It is instructive to compare the magnitude of Gini indices based on car prices with those available for Germany as a whole. There are several Gini indices available for different measures of welfare, as summarised in Table 2. The smallest Gini indices (0.289-0.299) are computed for net equivalence income of households, whilst the largest Gini indices (0.683-0.703) are based on wealth data. Our estimates presented in the two bottom rows of the table are similar in magnitude to the Gini index computed for market equivalence income, for which the corresponding figures for NUTS1 are not available.

Last but not least, in a pioneering attempt we apply our approach to estimate the Gini index for the NUTS2 regions. To the best of our knowledge, this has never been done for Germany before. Therefore, the outcome of this exercise, reported in Figure 7, defines another contribution of our paper to the literature.

4 Conclusion

We suggest an alternative indicator for measuring economic inequality among regions that is based on car sale prices placed in the Internet. Using Germany as an example we illustrate that our estimates of regional income levels as well as of Gini indices display high positive correlation with the official figures. In contrast to traditional measures, our data can be fast and inexpensively retrieved from the web, and, more importantly, allow estimating Gini indices at the NUTS2 level—something that has never been done for Germany before. We conclude that our approach to measuring regional inequality is a useful alternative source of information that could complement officially available measures.

References

- Askitas, N. and K. F. Zimmermann (2009). Google econometrics and unemployment forecasting. IZA Discussion Papers 4201, Institute for the Study of Labor (IZA).
- Becker, I. and R. Hauser (2003). Anatomie der Einkommenverteilung. Berlin.
- Choi, H. and H. Varian (2009). Predicting initial claims for unemployment benefits. Technical report, Google Inc.
- D'Amuri, F. and J. Marcucci (2009). "Google it!" Forecasting the US unemployment rate with a Google job search index. MPRA paper, University Library of Munich, Germany.
- Eurostat (2011). Regions in the European Union: Nomenclature of territorial units for statistics NUTS 2010/EU-27. Methodologies and working papers, Eurostat.

- Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant (2009). Detecting influenza epidemics using search engine query data. *Nature* 457, 1012–1014.
- Kholodilin, K. A., M. Podstawski, and B. Siliverstovs (2010). Do Google searches help in nowcasting private consumption? A real-time evidence for the US. KOF Working papers 10-256, KOF Swiss Economic Institute, ETH Zurich.
- Krause, P. and A. Schäfer (2005). Verteilung von Vermögen und Einkommen in Deutschland: Große Unterschiede nach Geschlecht und Alter. Wochenbericht nr. 11/2005, DIW Berlin.
- Sachverständigenrat (2011). Jahresgutachten 2011/12.
- Statistisches Bundesamt (2011). Gini-Koeffizient der Äquivalenzeinkommen. http://www.amtlichesozialberichterstattung.de/Tabellen/tabelleA3.html.

| Correlation between | | NUTS1 Mobile Autoscout24 Year | | NUTS2 Mobile Autoscout24 Year | | | |
|---|------------------------------|----------------------------------|-------|----------------------------------|-------|-------|------|
| | | | | | | | |
| Number of inhabitants | Number of ads | 0.995 | 0.998 | 2010 | 0.972 | 0.975 | 2010 |
| Gross Regional Product per capita, at current prices | Average car price | 0.725 | 0.843 | 2010 | 0.673 | 0.690 | 2008 |
| Gross national income per capita | Average car price | 0.830 | 0.885 | 2009 | • | | |
| Net national income (primary income) per capita | Average car price | 0.850 | 0.890 | 2009 | • | | |
| National income (wage and property income) per capita | Average car price | 0.858 | 0.884 | 2009 | • | | |
| Balance of primary income, net (uses) per capita | Average car price | 0.861 | 0.885 | 2009 | 0.794 | 0.694 | 2008 |
| Disposable income, net (uses) | Average car price | 0.805 | 0.901 | 2009 | 0.748 | 0.692 | 2008 |
| Gross wage per employee | Average car price | 0.801 | 0.815 | 2010 | 0.655 | 0.587 | 2009 |
| Private consumption per capita | Average car price | 0.772 | 0.897 | 2009 | | • | • |
| Gini index (net equivalence income) | Gini index (car-price based) | 0.809 | 0.768 | 2010 | | | |

Table 1: Correlation coefficients

Sources: All income measures are taken from the Arbeitskreis VGR der Länder, http://www.statistik.baden-wuerttemberg.de/Arbeitskreis_VGR/ergebnisse.asp#BIP_K. The Gini index is from the Statistisches Bundesamt (2011). The information on latest available official data is reported in the column Year.

| Welfare measure | | Gini index | |
|--|------|------------|--|
| | | | |
| Market equivalence income ¹ | 2009 | 0.486 | |
| Net equivalence income ¹ | 2009 | 0.289 | |
| Net equivalence $income^2$ | 2010 | 0.290 | |
| Net equivalence income ³ | 2002 | 0.299 | |
| $Gross wealth^4$ | 2002 | 0.703 | |
| Net wealth ⁴ | 2002 | 0.683 | |
| Based on car prices (mobile.de) ⁵ | 2011 | 0.429 | |
| Based on car prices $(autoscout24.de)^5$ | 2011 | 0.403 | |

Table 2: Gini indices for Germany

¹ Sachverständigenrat (2011), household income per person, in 2005 prices, with imputed rent.
 ² Statistisches Bundesamt (2011).
 ³ Krause and Schäfer (2005), household income per person, in 2002 prices, without imputed rent.
 ⁴ Krause and Schäfer (2005).
 ⁵ Authors' calculations.





Figure 2: Average car prices in 1000 Euro (mobile.de); NUTS1



Average car price





Figure 3: Average car prices (mobile.de) vs balance of primary income, per capita; NUTS2

Figure 4: Average car prices in 1000 Euro (mobile.de); NUTS2



Average car price



■ >16.4





Figure 6: Gini coefficient (based on car prices, mobile.de); NUTS1







Figure 7: Gini coefficient (based on car prices, mobile.de); NUTS2





