

Modelowanie wartości eksportu dóbr i usług

Justyna Klejdysz

Ekonometria bayesowska

28 stycznia 2017

Spis treści

1	Wstęp	2
2	Zbiór danych	2
3	Estymacja bayesowska	3
3.1	Postać modelu	3
3.2	Elicytacja rozkładu <i>a priori</i>	3
3.3	Wyniki estymacji i porównanie modeli	5
3.4	Prognoza	6
3.5	Bayesowskie łączenie wiedzy	7
4	Podsumowanie	9

1 Wstęp

Celem niniejszego opracowania jest zidentyfikowanie zmiennych, które wpływają na wartość eksportu dób i usług, oszacowanie modelu eksportu z wykorzystaniem metod bayesowskich w oparciu o dane dla 27 państw europejskich, a następnie wyznaczenie oczekiwanej wartości eksportu dla wybranych województw w 2014 r. Województwa są jednostkami spoza próby, dla których wartości eksportu nie są rejestrowane. Województwo jest rozpatrywane jako mała gospodarka otwarta. Eksport został zdefiniowany tutaj jako dostarczanie towarów i usług poza granice województwa - do innych województw lub poza granice kraju.

2 Zbiór danych

Zbiór danych podzielony jest na dwa podzbiory. Pierwszy podzbiór zawiera dane o 27 państwach europejskich w latach 2000 - 2014 r, łącznie 379 obserwacji. Panel nie jest zbilansowany. Drugi podzbiór, z takim samym zestawem zmiennych, dotyczy województw. Szczegółowy opis zmiennych i źródła danych przedstawia tabela 1. Wyboru zmiennych dokonano w oparciu o przegląd literatury, odzwierciedla on wiele podejść do objaśniania wielkości eksportu i w konsekwencji zaproponowany zestaw zmiennych jest relatywnie szeroki w porównaniu do większości opracowań poświęconych modelowaniu potencjału eksportowego w ujęciu makroekonomicznym. W części pracy poświęconej bayesowskiemu łączeniu wiedzy podjęto próbę odpowiedzi na pytanie, które z zaproponowanych regresorów powinny być uwzględnione i jaka jest ich waga w wyjaśnianiu wartości eksportu.

Tabela 1: Opis zmiennych i źródła danych

<i>Zmienna</i>	<i>Opis</i>	<i>Źródło</i>
logexpgoodsserv	Eksport towarów i usług (mln eur, CLV*, logarytm)	Eurostat
reer37lc	Realny efektywny kurs walutowy (deflator: jednostkowe koszty pracy); dla województw: CPI	Eurostat, BDL
logvaperempl	Wartość dodana brutto na pracownika (tys. eur, CLV, logarytm)	Eurostat
loggdpw	Ważona suma PKB 15 najważniejszych partnerów handlowych (mln dol., ceny stałe 2010, logarytm)	COMTRADE, WDI, obliczenia własne
logpoptot	Liczba ludności (tys., logarytm)	Eurostat
edu_secter	Odsetek ludności w wieku 25-64 z wykształceniem średnim lub wyższym	Eurostat
roadssurf	Gęstość dróg na 100 km ²	Eurostat, WDI, BDL
prodind	Dynamika produkcji przemysłowej	Eurostat, BDL
emplagr_frac	Odsetek zatrudnionych w rolnictwie	Eurostat, BDL
eu	Zmienna binarna: 1 - państwo w UE	

* CLV - *Chain Linked Volumes*

3 Estymacja bayesowska

3.1 Postać modelu

Przedmiotem analizy będzie model z efektami indywidualnymi α_i dla państw, z niehierarchicznym rozkładem *a priori*:

$$\mathbf{y}_i = \alpha_i \boldsymbol{\iota}_T + \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i = \mathbf{X}_i^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon}_i \quad (1)$$

$$\mathbf{X}^* = \begin{bmatrix} \boldsymbol{\iota}_T & \mathbf{0}_T & . & . & \mathbf{X}_1 \\ \mathbf{0}_T & \boldsymbol{\iota}_T & \mathbf{0}_T & . & \mathbf{X}_2 \\ . & \mathbf{0}_T & . & . & . \\ . & . & . & . & . \\ \mathbf{0}_T & . & . & \boldsymbol{\iota}_T & \mathbf{X}_N \end{bmatrix} \quad \boldsymbol{\beta}^* = \begin{bmatrix} \alpha_1 \\ . \\ . \\ \alpha_N \\ \boldsymbol{\beta} \end{bmatrix} \quad \varepsilon_{it} \sim N(0, h^{-1}) \quad i.i.d. \quad (2)$$

gdzie \mathbf{X}^* jest macierzą o wymiarach $NT \times (N+k)$, k to liczba zmiennych, nieuwzględniająca stałej. Przyjęto rozkład Normalny-Gamma dla parametrów modelu:

$$\boldsymbol{\beta}^* \sim N(\underline{\boldsymbol{\beta}}, \underline{\mathbf{U}}) \quad (3)$$

$$h \sim G(\underline{s}^{-2}, \underline{v}) \quad (4)$$

3.2 Elicytacja rozkładu *a priori*

Przeprowadzono następującą analizę, mającą na celu wyznaczenie rozkładów *a priori*: $\underline{\boldsymbol{\beta}}$, $\underline{\mathbf{U}}$, \underline{s} oraz \underline{v} .

- Aprecjację realnego efektywnego kursu walutowego interpretuje się najczęściej jako utratę konkurencyjności na rynkach międzynarodowych. Oczekuje się, że wpływ zmiennej na wartość eksportu będzie niedodatni. Podsumowanie poprzednich analiz przeprowadzanych dla państw strefy euro i UE (Allard 2009, Sertić et al. 2015, Ca' Zorzi et al. 2007) wskazuje na zakres elastyczności eksportu względem realnego kursu walutowego od około -0.58% do ok. -0.01%. Na tej podstawie wartość oczekiwaną *a priori* ustalono na -0.003 (indeks nie został zlogarytmowany). Niepewność co do siły wpływu wyraża wariancja *a priori* ustalona na 10^{-5} . Oznacza to przypisanie 95% prawdopodobieństwa, że wartość współczynnika znajduje się w przedziale -0.009 do 0.003, który obejmuje zakres wartości przywoływanych w literaturze.
- Dokonując syntezy wyników otrzymanych dla strefy euro, Unii Europejskiej i OECD,

ustalono wartość oczekiwaną *a priori* dla parametru stojącego przy zmiennej dotyczącej wielkości popytu zewnętrznego (logarytm ważonej sumy PKB 15 największych partnerów handlowych) na poziomie 0.8. Współczynnik jest z pewnością większy od 0 i na podstawie przeglądu literatury jego wariancję *a priori* ustalono na 0.16, przypisując 95% prawdopodobieństwo, że wartość elastyczności znajduje się w przedziale od ok. 0.02 do 1.6. Długookresowa elastyczność wielkości eksportu względem popytu światowego mniejsza od 1 oznaczałaby skłonność do utraty przez kraje Unii Europejskiej udziałów w eksporcie w miarę wzrostu popytu światowego (Krugman, 1989), ale może być też związana ze zwiększaniem zagranicznych inwestycji bezpośrednich, które substytuują krajowy eksport (ECB, 2005).

- UNCTAD (2004) szacuje, że wzrost populacji o 1% przekłada się przeciętnie na wzrost wartości eksportu od 0.002% do 0.241% wśród państw wysoko rozwiniętych. Na tej podstawie ustalono wartość oczekiwaną *a priori* na poziomie 0.12, z wariancją 0.03, co oznacza, że przypisano 95% prawdopodobieństwo, że jego wartość znajduje się w przedziale od ok. -0.2 do 0.45.
- Wartość oczekiwaną *a priori* dla parametru wyrażającego siłę wpływu produktywności pracy na wartość eksportu ustalono jako 1, co wyraża przypuszczenie, że wzrost produktywności o 1% przekłada się na około 1% wzrost wartości eksportu. Jednocześnie ustalono wysoką wariancję *a priori* na poziomie 4, ponieważ brakuje analiz panelowych dla państw europejskich, które w równaniach eksportu uwzględniałyby produktywność pracy, mierzoną wartością dodaną brutto na pracownika. Jednak na podstawie badań przeprowadzonych dla pojedynczych państw lub uwzględniających w zestawie regresorów PKB, można oczekiwać silnego dodatniego wpływu tej zmiennej.
- Przypuszcza się również, że wyższy odsetek ludności z wykształceniem średnim lub wyższym pozytywnie oddziałuje na wielkość eksportu, podobnie gęstość sieci drogowej, wskaźnik produkcji przemysłowej i przynależność do UE, jednak trudno precyzyjnie oszacować siłę wpływu powyższych zmiennych. Ustalono wartości oczekiwane odpowiadające punktowym oszacowaniom uzyskanym w wyniku przeprowadzenia innych metod estymacji: FM-OLS, FE, ale na tym samym zbiorze danych. Dlatego przyjęto wysokie wariancje *a priori*.
- Nie ma podstaw, aby przypisywać wysokie prawdopodobieństwo określonym wartościom efektów indywidualnych dla państw, dlatego ustalono dla nich nieinformacyjny rozkład *a priori*.
- Model wyraża zależności pomiędzy zmiennymi makroekonomicznymi, dlatego oczekuje się, że wariancja wartości eksportu jest objaśniana w dużym stopniu, a wariancja reszt jest niska. Odchylenie standardowe składnika losowego ustalono na 0.22 (około

3-krotnie więcej niż wynika z klasycznej estymacji modelu z efektami stałymi). Liczbę stopni swobody *a priori* ustalono na 100, co zwiększa liczbę stopni swobody o ponad 25% w porównaniu do wyjściowego zbioru. Relatywnie duża waga wiedzy *a priori* wynika z jednej strony z dość krótkiej próby (uwarunkowanej ograniczoną dostępnością kompletnych danych), z drugiej strony z potwierdzonych w licznych opracowaniach zależności makroekonomicznych. Wartości oczekiwane i odchylenia standardowe *a priori* podsumowuje tabela 2.

3.3 Wyniki estymacji i porównanie modeli

Znaki oszacowanych średnich *a posteriori* są zgodne z oczekiwaniami. Odchylenia standardowe są relatywnie wysokie w stosunku do średnich, ale niższe niż w przypadku nieinformacyjnego rozkładu *a priori* (por. tabela 2). Informacje o HPDI i ilorazie szans *a posteriori* zestawione w tabeli 3 pozwalają wstępnie ocenić, czy pojedyncze współczynniki są różne od zera i jakie jest prawdopodobieństwo modelu z restrycją nierównościową nałożoną na pojedynczy parametr w porównaniu do prawdopodobieństwa modelu bez restrykcji.

Tabela 2: Wartości oczekiwane rozkładów *a priori* i *a posteriori* (skrajnie nieinformacyjnego i informacyjnego). W nawiasach podano błędy standardowe. W tabeli pominięto oszacowania dla efektów indywidualnych.

	<i>a priori</i>	<i>a posteriori</i>	
	informacyjny	skrajnie nieinformacyjny	informacyjny
logvaperempl	1.000 (2.236)	2.050 (1.575)	1.621 (0.993)
reer37lc	-0.003 (0.003)	0.001 (0.008)	-0.002 (0.003)
loggdpw	0.800 (0.400)	0.099 (0.555)	0.248 (0.198)
logpoptot	0.120 (0.173)	0.050 (2.492)	0.095 (0.171)
roadssurf	0.003 (0.707)	0.002 (0.006)	0.002 (0.005)
prodind	0.000 (0.707)	0.001 (0.012)	0.000 (0.009)
edu_secter	0.010 (0.707)	0.012 (0.024)	0.012 (0.016)
emplagr_frac	0.010 (0.707)	0.017 (0.078)	-0.010 (0.060)
eu	0.030 (0.707)	0.034 (0.322)	0.103 (0.248)

Po uwzględnieniu wiedzy *a priori* oszacowany wpływ *a posteriori* dla realnego efektywnego kursu walutowego jest ujemny - zgodnie z teorią oznacza to, że wzrost realnego efektywnego kursu walutowego wiąże się z obniżeniem konkurencyjności eksportu. Zarówno 95%, jak i 99% przedział HPDI dla współczynnika stojącego przy tej zmiennej nie zawiera 0. Klasyczne podejście do estymacji wskazuje z kolei na jej nieistotność, a oszacowanie parametru jest większe od zera. Przedziały ufności HPDI wskazywałyby, że jedynie współczynniki dla indeksu produkcji przemysłowej i zatrudnienia w rolnictwie są bliskie 0.

Dla większości współczynników iloraz szans *a posteriori* pozwala stwierdzić tylko nieznaczną przewagę modelu z restrykcją, odpowiadającą wiedzy *a priori* w porównaniu do modelu, gdy ta restrykcja nie jest spełniona. Jedynie model nakładający restrykcję w postaci dodatniego wpływu wartości dodanej brutto na pracownika jest zdecydowanie bardziej prawdopodobny niż model bez restrykcji. W przypadku modelu z restrykcją nałożoną na współczynnik przy zmiennej wskazującej na wielkość popytu zewnętrznego prawdopodobieństwo jest ok 50% większe niż modelu niespełniającego restrykcji. Prawdopodobieństwo modelu, w którym na wszystkie współczynniki nałożono restrykcje zgodne z wiedzą *a priori* wynosi 5%.

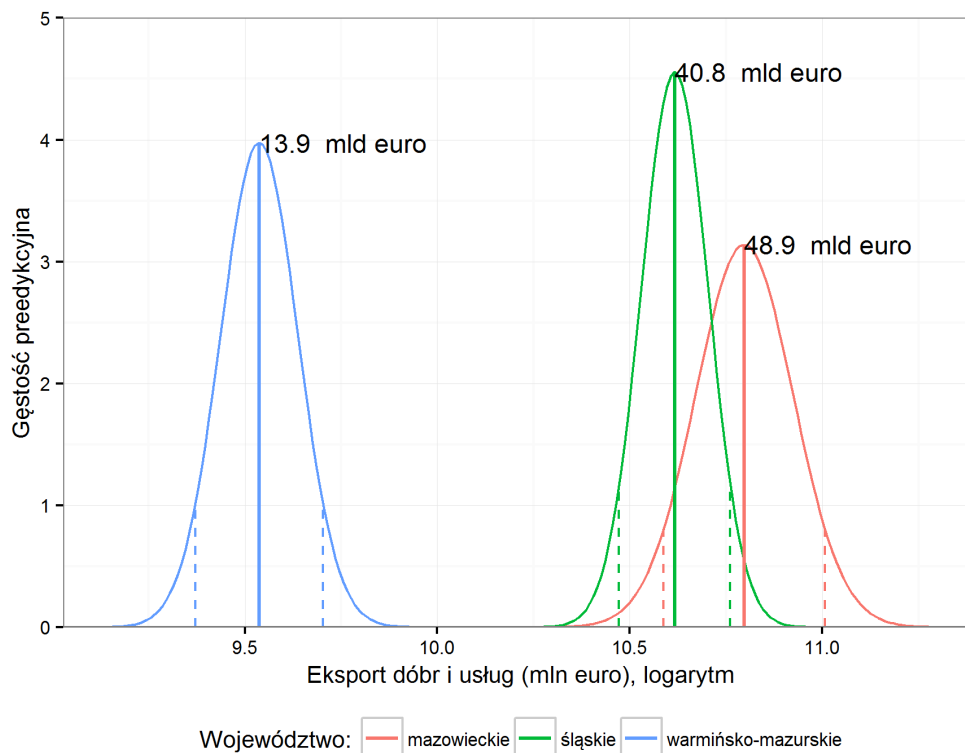
Tabela 3: Przedziały ufności HPDI i iloraz szans *a posteriori* dla modelu z restrykcją w postaci nierówności

	Granice 95% HPDI		Granice 99% HPDI		PO $\beta_j > 0 y$
	dolna	górną	dolna	górną	
logvaperempl	1.333	1.908	1.242	1.999	18.027
reer37lc	-0.003	-0.001	-0.003	-0.001	0.997
loggdpw	0.191	0.306	0.173	0.324	1.488
logpoptot	0.046	0.145	0.030	0.161	1.165
roadssurf	0.001	0.003	0.000	0.004	1.003
prodind	-0.003	0.003	-0.003	0.004	1.000
edu_sector	0.007	0.017	0.006	0.018	1.019
emplagr_frac	-0.028	0.007	-0.034	0.013	0.983
eu	0.031	0.175	0.008	0.197	1.178

3.4 Prognoza

Prognozę eksportu przeprowadzono dla trzech województw: mazowieckiego, śląskiego i warmińsko - mazurskiego. Jako efekt indywidualny dla każdego z województw przyjęto średnią oszacowanych wartości oczekiwanych efektów indywidualnych dla państw. Obserwowane wartości zmiennych objaśniających wykorzystano do oszacowania parametrów trzech rozkładów brzegowych wielowymiarowego rozkładu predykcyjnego t . Rozkłady predykcyjne, ich wartości oczekiwane oraz 90% HPDI przedstawia rysunek 1.

Rysunek 1: Rozkłady predykcyjne wartości eksportu dóbr i usług dla województwa mazowieckiego, śląskiego i warmińsko-mazurskiego. Linia przerywaną zaznaczono 90% HPDI



3.5 Bayesowskie łączenie wiedzy

Dotychczasowa analiza wykazała, że niektóre spośród wybranych zmiennych mogą mieć małą wagę w objaśnianiu wielkości eksportu. Dysponując dziewięcioma zmiennymi makroekonomicznymi, istnieje ponad 500 możliwych specyfikacji modelu. Bayesowskie łączenie wiedzy umożliwia uzyskanie parametrów rozkładu *a posteriori*, będących średnią ważoną wyników możliwych modeli, z wagami będącymi prawdopodobieństwami modelu *a posteriori*:

$$p(\theta|y) = \sum_{r=1}^R p(\theta|y, M_r)p(M_r|y) \quad (5)$$

Konieczny jest wybór prawdopodobieństwa modelu *a priori*. Zastosowano dwa podejścia do jego ustalenia. W pierwszym przyjęto, że rozkład prawdopodobieństwa modelu *a priori* jest rozkładem dwumianowym, w którym każda zmienna może być niezależnie włączona do modelu z prawdopodobieństwem 1/2. Wówczas oczekiwana wielkość modelu wyniesie 4 lub 5 (por. rysunek 2a). W drugim podejściu arbitralnie określono prawdopodobieństwa włączenia każdej zmiennej, w oparciu o przegląd literatury. Zmiennym produktywność pracy, realny efektywny kurs walutowy i popyt zewnętrzny przypisano wyższe prawdopodobieństwo

włączenia do modelu: 0.8, dla pozostałych zmiennych pozostawiono 0.5 (por. rysunek 2b). Drugie podejście skutkuje zwiększeniem oczekiwanej wielkości modelu *a posteriori*.

Dla parametru precyzji h standardowo przyjmuje się nieinformacyjny rozkład *a priori*. Należy również wybrać parametr *g-prior* Zellnera, który określi postać rozkładu *a priori* dla współczynników β_r :

$$\beta_r|h \sim N\left(\mathbf{0}_{k_r}, h^{-1}\left(\frac{1}{g_r}X_r'X_r\right)^{-1}\right) \quad (6)$$

Wysoka wartość *g-prior* wyraża niepewność, że współczynniki są równe 0 (Zeugner 2011). Najczęściej ustala się $g\text{-prior} = N$. Im jego wartość jest wyższa, tym oszacowania będą w większym stopniu uwzględniały informację zawartą w danych.

Zamiast generowania zmiennych binarnych dla efektów indywidualnych, zastosowano transformację *within*, odejmując od wartości zmiennych średnie wartości dla danej jednostki panelu. Prawdopodobieństwa *a posteriori* włączenia poszczególnych zmiennych do modelu oraz oszacowania parametrów *a posteriori* przedstawiono w tabeli 4.

Rysunek 2: Rozkład wielkości modelu *a posteriori* w zależności od wyboru rozkładu *a priori*: (a) prawdopodobieństwo *a priori* włączenia zmiennej: 1/2, (b) indywidualne określenie prawdopodobieństw *a priori* włączenia zmiennej.

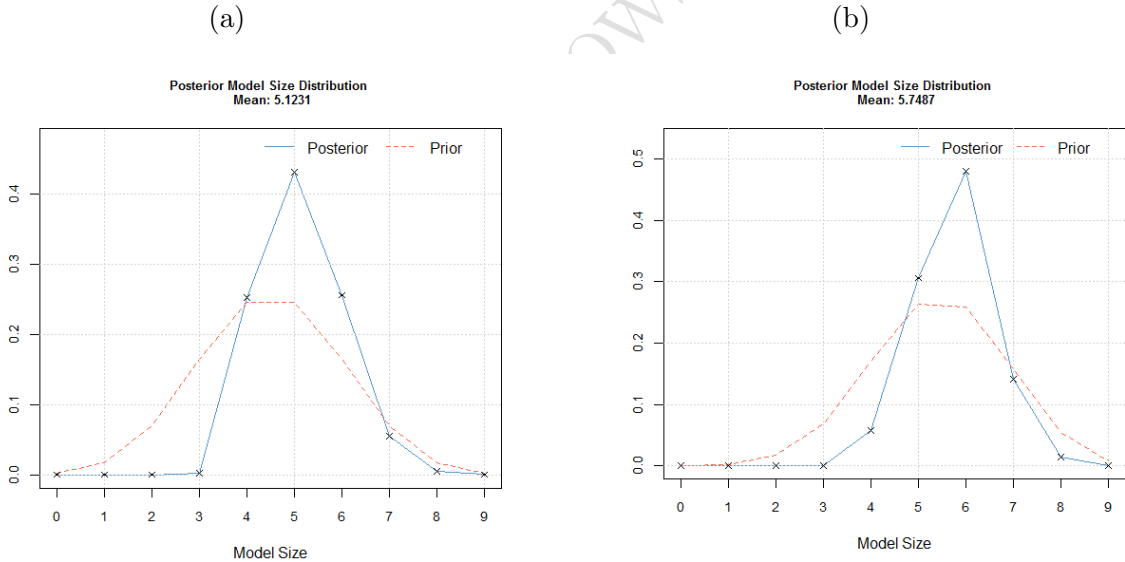


Tabela 4: Bayesowskie łączenie wiedzy - wyniki

	P-stwo <i>a priori</i> modelu (a)			P-stwo <i>a priori</i> modelu (b)		
	Post. Prob.	Post Mean	Post SD	Post. Prob.	Post Mean	Post SD
logvaperempl	1.000	2.146	0.102	1.000	2.125	0.106
edu_secter	1.000	0.012	0.001	1.000	0.012	0.001
roadssurf	1.000	0.002	0.000	1.000	0.002	0.000
emplagr_frac	0.972	0.018	0.006	0.962	0.017	0.006
loggdpw	0.599	0.056	0.053	0.881	0.085	0.046
reer37lc	0.231	0.000	0.000	0.586	0.001	0.001
eu	0.160	0.006	0.015	0.099	0.000	0.000
prodind	0.081	0.000	0.000	0.062	0.007	0.050
logpoptot	0.081	0.013	0.065	0.158	0.005	0.015

Wartości oszacowanych odchyłeń standardowych są na ogół niskie w porównaniu do wartości średnich *a posteriori*. Pozwala na większą pewność przy określeniu, które zmienne są ważne w objaśnianiu wielkości eksportu niż w przypadku analizy zaprezentowanej w części 3.3. Należałoby rozważyć, czy zasadne jest uwzględnianie dodatkowo zmiennej wskazującej na przynależność do UE, wskaźnika produkcji przemysłowej i liczby ludności, ze względu na niskie prawdopodobieństwa *a posteriori* włączenia do modelu. Sprawdzono również, że wspólnym zestawem zmiennych dla trzech najlepszych modeli jest wartość dodana brutto na pracownika, odsetek osób z wykształceniem średnim lub wyższym w wieku 25-64, gęstość dróg, odsetek zatrudnionych w rolnictwie i popyt zewnętrzny. Model o najwyższym prawdopodobieństwie *a posteriori* (wynoszącym 0.36) zawiera dodatkowo zmienną dotyczącą realnego efektywnego kursu walutowego. W preferowanym modelu znalazłoby się zatem 5 zmiennych.

4 Podsumowanie

W niniejszej pracy przeprowadzono krótką analizę na podstawie niehierarchicznego modelu regresji panelowej z efektami indywidualnymi. Zastosowano szereg upraszczających założeń o postaci rozkładów *a priori* i rozkładzie danych, które umożliwiły analityczne wyznaczenie rozkładu parametrów *a posteriori* oraz rozkładu predykcyjnego eksportu dla województw. Rozszerzenia obejmują m. in. skonstruowanie modelu hierarchicznego i uwzględnienie wymiaru czasowego panelu, szersze wykorzystanie metod numerycznych do porównywania modeli lub wyznaczenie prognozy w oparciu o model uwzględniający restrykcje w postaci nierówności dla niektórych zmiennych jako informację *a priori* i uwzględnienie efektów przestrzennych.

Bibliografia

- Allard, C. 2009. *Competitiveness in Central-Europe; What Has Happened Since EU Accession?* IMF Working Papers 09/121. International Monetary Fund.
- Ca' Zorzi, Michele, and Bernd Schnatz. 2007. *Explaining and forecasting euro area exports: which competitiveness indicator performs best?* Working Paper Series 0833. European Central Bank.
- Fugazza, Marco. 2004. "Export performance and its determinants: supply and demand constraints."
- Koop, G. 2003. "Bayesian Econometrics." *Wiley*.
- Mauro, Filippo di, Robert Anderton, Ekkehard Ernst, Remy Lecat, Mark Cassidy, Roberto Tedeschi, Erik Walch, Jurriaan Eggelte, Karin Wagner, Tobias Sebastian Blattner, et al. 2005. "Competitiveness and the export performance of the euro area."
- Sertić, Martina Basarac, Valentina Vučković, and Blanka Skrabić Perić. 2015. "Determinants of manufacturing industry exports in European Union member states: a panel data analysis." *Economic Research-Ekonomska Istraživanja* 28 (1): 384–397.
- Zeugner, Stefan. 2011. "Bayesian model averaging with BMS." *Tutorial to the R-package BMS 1e30*.