National Research University Higher School of Economics

as a manuscript

Galina Besstremyannaya

Econometric models for analyzing heterogeneity of economic agents

Summary
Dissertation for the purpose of obtaining academic degree
Doctor of Science in Economics

Academic adviser:
Professor, Ph.D. in Economics
Ruben Enikolopov

JEL: C21, C23, C44, D22, D24, G21, I13, I18

Moscow – 2023

1 Motivation

1.1 In quest of heterogeneity

The emergence of econometrics may be dated to the 19th century which saw a particular interest in the study of mean tendencies through statistical and regression analysis (Galton and Dickson, 1886, Galton, 1886, Quetelet, 1842, Gauss, 1823).[1] In those days, scientists were fascinated by their ability to establish statistical regularities for an average individual. In 1890, Sir Arthur Conan Doyle makes his hero, Mr. Sherlock Holmes, praise the merits of the approach: "[W]hile the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty. You can, for example, never foretell what any one man will do, but you can say with precision what an average number will be up to." (Conan Doyle (1890), The Sign of Four, P.196).

Originally, econometrics used least squares methods, since they "provide a general approach to estimating conditional mean functions" (Koenker (2005), P.1). It has taken economists long to admit the need of going beyond the analysis of a mean or median tendency in order to focus on differences across agents. In 1975, the manuscript by G.W.Bassett and R.Koenker on conditional quantiles of the dependent variable was rejected by *Econometrica* and by *Annals of Statistics* as reviewers feared the lack of scientific importance of the topic: "It may be of interest to compute regression analyses to minimize the sum of absolute deviations between the observed and fitted responses... But why should one consider $\tau \neq 0.5$?" (Bassett and Koenker (2017), P.4).

A theoretical recognition of "diversity in motivations" and "deep-seated heterogeneity of the subject matter of economics" (Sen (2004), P.583) became a well-established paradigm only by the end of the 20th century. The period also witnessed "empirical discoveries ... on the pervasiveness of heterogeneity and diversity in economic life" (Heckman (2001), P.674). They were brought about by the development of microeconomic theory in the first half of the 20th century, the collection of large datasets on consumers and producers in the second half of the century, and the expansion of statistical methods and computational means for the applied analysis (Heckman, 2001).

1.2 Definitions and econometric models of agent heterogeneity

Macro and micro economists interpret heterogeneity as the facts that economic agents differ in their economic, social, psychological, anthropological and other characteristics, and that these differences impact agent decisions (Heathcote et al., 2009, Blundell and Stoker, 2007, Browning et al., 1999). For instance, there exists "heterogeneity in individual tastes, heterogeneity in income and wealth risks, and heterogeneity in market participation" (Blundell and Stoker, 2007, P.4610).

In statistical and econometric terms, heterogeneity may be defined as "information about [relevant variables] known to agents and acted on in their choices" (Cunha et al. (2005), P.3). Heterogeneity is reflected in "the dispersion in factors that are relevant *and known* to individual agents when making a particular decision" (Browning and Carro (2007), P.47, italicized in the original).

Econometricians distinguish observed and unobserved heterogeneity. *Observed* heterogeneity is

[1]See historic reviews in Koenker (2017), Angrist and Pischke (2009), Stigler (1997).

commonly dealt with through inclusion of a detailed list of agent's characteristics in the list of explanatory variables in regression. The productivity analysis offers additional interpretation of heterogeneity: the impact of agent characteristics and of the so-called environmental variables[2] on efficiency scores (Fried et al., 2008, Coelli et al., 2005, Simar and Wilson, 2008).

Unobserved heterogeneity is present when "those relevant factors ... are known to the agent but not to the researcher", Browning and Carro (2007), P.48. The most prevalent approach of incorporating this type of heterogeneity in the applied econometric analysis till the early 2000s was limited to consideration of the panel data fixed effect models: the fixed effects (individual effects) were regarded as a reflection of individual-specific or firm-specific unobserved heterogeneity.[3] Another widespread method of accounting for unobserved heterogeneity was the use of instrumental variable techniques aimed at overcoming the omitted variable bias.

The recognition of differences in preferences across consumers and producers led to a new interpretation of heterogeneity: heterogeneity implies that the effect of agent characteristics on economic choices may be different across *groups* of agents (Heathcote et al., 2009, Browning et al., 1999).

Plausible expectations about the existence of such heterogeneity are founded on the narrowness of a pure "economic approach to human behaviors" (Sen (2004), P.604). Indeed, various norms which are specific to social and peer groups lead to selection among different types of motivations (Brock and Durlauf, 2001, Sen, 2004). Moreover, emerging experimental and empirical literature causes researchers to cast doubts about the validity of rationality assumption, to reconsider the internal consistency of agent choices and incorporate altruistic behavior, as well as tendency to experiment, adapt and expect, in the decision-making by individuals and firms (Browning and Carro, 2007, Kirman, 2006, Cunha et al., 2005, Sen, 2004).

But the awareness of econometricians about this type of heterogeneity, i.e. "heterogeneity of a different sort, associated with the [different] coefficient vectors" [for different subsamples of observations] (Greene (2003), P.359) – began to be gradually observed in various fields of applied economics only in the 2000s-2010s. The rapidly developing application of econometric methods include finite mixture (latent class models), conditional quantile regression, conditional average treatment effects, and dynamic panel data models (Schennach, 2020, Angrist and Pischke, 2015, Cameron and Trivedi, 2013, Greene, 2012, Wooldridge, 2011, Angrist and Pischke, 2009, Chernozhukov and Hansen, 2008).

Latent class models, as is further described in the brief literature review section in this summary, is a classic example of dealing with unobserved heterogeneity and groupwise differences: classes are latent and probabilistic. Other methods (e.g., conditional quantile regression and conditional average treatment effect) may be taken as the means to account for observed heterogeneity across groups of agents.

1.3 Research agenda

As was noted by James Heckman in his Nobel lecture given at the turn of the 21st century, the essential tasks of modern microeconometrics are "to unite theory and evidence and to evaluate policy interventions" (Heckman (2001), P.673). Regarding microeconometric evidence that requires identification

[2]Macroeconomic variables or firm-level variables which are not directly controlled by producers.

[3]See Verbeek (2004), P.353, Hayashi (2000), P.325, Wooldridge (2012), P.456, Greene (2003), P.310, Baltagi (2005), PP.14–15, 19, 135–136.

of heterogeneity between agents and groups of agents, econometricians have reached a general understanding about the existence of "differences in the variances of the disturbances across groups" (Greene (2003), P.546) and an agreement that "heterogeneity across groups ... is typical in microeconomic data" (ibid, P.359). However, inadequate attention is still given to the analysis of heterogeneity of economic agents and of heterogeneous effects of policy reforms (Angrist and Pischke, 2009, Browning and Carro, 2007, Kirman, 2006, Sen, 2004). Even the question on whether tastes differ across individuals, which was posed in the seminal and provocative paper by Stigler and Becker (1977), required a continuation of the discussion in the 2000s (Brock and Durlauf, 2001).

In fact, a wide range of questions raised by microeconomists supports the cause for empirical identification of observed and unobserved heterogeneity across economic agents in general and groups of agents (e.g. subpopulations of individuals and firms) in particular, and urge quantification of the heterogeneous impact of policy reforms on these agents. Examples of such questions are listed below.

1. Why are there productivity differences across firms? What is the interrelation between differences in management, productivity and firm growth? What issues related to local markets and overall macroeconomic environment explain inefficiency of firms (which is commonly defined as deviation from the production possibility frontier or from cost-minimization trajectory)?[4] Do economies of scale and scope differ across low-cost and high-cost firms? Does elasticity of output with respect to labor, capital and materials differ at high-output and low-output firms of a given industry?
2. If there are differences in firm productivity or firm costs, and in their time profiles, what are the consequences for the policy-makers? Specifically, what are differential effects of policy regulation on more/less productive firms or firms with higher/lower costs?
3. Can demand by groups of consumers respond differently to changes in the price of the product? Are there consumers with inelastic demand? If this is the case, are there any (hidden) inequities in consumer demand which need to be incorporated in welfare analysis and policy regulation?
4. Do firms in publicly regulated industries respond differently to price or quality contracts induced by the social-planner? Can the same regulation positively impact the performance of some firms but have a negative effect on the performance of others? What are the causes and consequences of such a heterogeneous response to policy reforms?

2 Objectives of the research

2.1 Outline of objectives

The purpose of this research is to develop econometric models in order to reveal heterogeneity in economic choices by producers and consumers, to disentangle heterogeneous effects of exogenous shocks on firm costs, and to evaluate heterogeneous effects of policy reforms aimed at price and quality regulation.

[4] Indeed, within the production possibility set in each industry, such commonly unobserved variables as poor management or lack of knowledge about the applicability of technology to production at a given firm may lead to productive/cost inefficiencies (Bloom et al., 2016, Bloom and van Reenen, 2010, Griliches, 1996). See numerous reviews on efficiency and productivity analysis, e.g. (Tone, 2017, Fried et al., 2008, Coelli et al., 2005).

The *theoretical* part of the research has the following objectives.

1. To develop a methodology for the bias-correction of the data envelopment scores in the cost-minimization problems of Färe et al. (1985) and Tone (2002).
2. To investigate the applicability of the conditional quantile regression estimator with quantile-independent fixed effects (Canay, 2011) in cases of short panels.
3. To study the means of correcting the asymptotic bias of the conditional quantile regression estimator in the short panels with quantile-dependent and quantile-independent fixed effects.
4. To account for multivariate dependence of the policy variable in dynamic panel data models and disentangle two sources of intertemporal dependence: the policy effect and the impact of regression towards the mean.

The *empirical* part of the research has both *economic* and *econometric* objectives. The economic objectives are listed below. For the sake of brevity, the below list omits the repetition of the fact that each *economic* objective required a development/modification of an *econometric* model in order to account for observed and/or unobserved heterogeneity of agents.

Each objective is formulated and analyzed in a way which pertains to a general setting within microeconometrics of productivity analysis, regulation, contract theory, or policy evaluation. At the same time, each empirical application and econometric model deal with the data for firms in a particular industry, with consumer demand for certain goods and services, as well as with examples of price or quality regulation, targeted at producers or consumers.

1. To explore the relationship between management and cost efficiency of public enterprises, as well as the time profiles of cost efficiency.
2. To estimate the conditional average treatment effect of a reform aimed at stimulating yardstick competition in public enterprises (the so-called prospective payment system which gives a fixed reimbursement for each type of product, regardless of the actual costs of production) on technical and cost efficiency of the enterprises.
3. To disentangle the differential effect of declining rates in the prospective payment system on the output and quality of public enterprises.
4. To evaluate the differential impact of the introduction of the intertemporal incentive contract on the performance of economic agents with different values of the pre-reform performance.
5. To reveal the behavioral differences in estimating consumer demand for a "necessity good" and to identify price elastic and price inelastic subpopulations of consumers.
6. To measure the heterogeneous treatment effect of price changes on consumption of a "necessity good".
7. To assess equity of access to a "necessity good" by consumers with high and low need of this good.
8. To reveal the heterogeneous effect of macroeconomic shocks on the time profiles of costs at high-cost and low-cost financial institutions as well as to discover differences in their economies of scale and scope.
9. To study the impact of different forms of regional social institutions on quality of public goods, using the example of healthcare provision and institutional environment which allows private health insurers to operate within the mandatory health insurance system.

10. To disentangle the differences in the productivity of capital and labor, and to evaluate the optimality of the labor/capital mix at high-output and low-output public enterprises.
11. To evaluate the differences in the association between R&D-to-sales ratio and firm growth at fast-growing and slow-growing high-tech innovative firms.

The research tasks 1–3, in the empirical group of tasks, are applied to the evaluation of cost and technical efficiency of Japanese acute-care local public hospitals in the early 2000s, while task 4 concerns the analysis of quality of the US acute-care Medicare hospitals in the 2010s. Task 5 in the empirical group, as well as task 1 in the theoretical group, deal with the Japanese banks in the 2000s-early 2010s. Tasks 6–8 are applied to the study of consumer demand for medical care in Japan. Task 9 deals with the analysis of Russian regions in 2000s-2010s and investigates the impact of private health insurers on the quality of regional healthcare systems. Task 10 examines the productivity of Japanese acute-care local public hospitals over the past two decades, while task 11 concerns the analysis of the growth of Japanese high-tech manufacturing firms in the 2010s.

The empirical part of the research uses macro-level and micro-level data for the US, Japan and Russia. One group of datasets are microdata on nationwide samples of Japanese firms, banks and acute-care local public hospitals (Orbis, Bankscope, Nikkei NEEDs, Yearbooks of Local Public Enterprises, Financial Statements of Banks). Another group are microdata on the nationwide samples of the US acute-care Medicare hospitals by Centers for Medicare and Medicaid. The third group are consumer-level data for Japan (representative surveys: Japan Panel Survey of Consumers and Keio Household Panel Survey) [5] and for the US (extracts from the census). Macro-data include country-level variables from international organizations (the OECD, IMF, WHO), national ministries and statistical agencies (e.g. Bank of Japan, Statistical Bureau of Japan, Federal State Statistics Survey (Rosstat), Russian Ministry of Finance, website "Insurance in Russia").

The research tasks are analyzed in the main part of this dissertation in 20 articles in the Web of Science/Scopus journals. These are 14 physical articles (13 of them written in English) of which 7 are sole-authored, 4 sole-authored articles and 2 co-authored articles are in the A list of the HSE journals (4 sole-authored articles and 1 co-authored article are also in the Scopus/WoS journals of the first quartile). Under the double weight of the HSE A list and/or WoS/Scopus first quartile journals, the number of articles becomes 20. The articles deal with theoretical and applied issues of econometric modeling.

Extensions of the dissertation are available in the Supplement to this summary. They are published in 5 additional sole-authored articles (2 of the WoS/Scopus journals in economics, 1 in the HSE D list of peer-reviewed economics journals, 2 of the WoS/Scopus journals in development or management) and provide statistical, economic and econometric insights into heterogeneous behavior of economic agents. Other extensions are available in peer-reviewed REPEC publications.

2.2 Identification of heterogeneity

Several approaches to the identification of observed and unobserved agent heterogeneity are used in the dissertation in order to achieve the objectives of the research. The list below outlines the approaches, the

[5] The cooperation of The Keio University Panel Data Research Center (Tokyo) and of The Institute for Research on Household Economics (Tokyo) for respectively providing the data of the Japan Household Panel Survey and of the Japanese Panel Survey of Consumers is gratefully acknowledged.

models, and gives examples of corresponding papers.

Observed heterogeneity

Observed heterogeneity may be defined as the fact that a certain variable is relevant for an agent's decision and there is variance in the values of the variable across agents, see Browning and Carro (2007) and Cunha et al. (2005).

The basic approach for identification of observed heterogeneity implies the inclusion of a covariate in the regression. The significance of the estimated coefficient for the covariate implies agent heterogeneity in view of the impact of this covariate on the dependent variable. The implementation of the approach is conducted through the following econometric models:

1. an OLS or a non-parametric regression as the main model (Besstremyannaya, 2015b, 2009a),
2. post-estimation analysis in productivity research, with OLS regression applied to SFA/DEA efficiency score (Besstremyannaya, 2013, Besstremyannaya and Simm, 2019) or to the residual in the conditional quantile regression (Besstremyannaya, 2017a, Besstremyannaya and Golovan, 2022b, Besstremyannaya et al., 2022),
3. post-estimation analysis in policy evaluation and/or measuring the average treatment effect/conditional average treatment effect (Besstremyannaya, 2015a).

Advanced approaches are targeted at the identification of groupwise heterogeneity, i.e. at finding statistical differences in the estimated coefficients for the covariate at groups of observations. The econometric models below are employed in the dissertation for this purpose:

1. a conditional quantile regression (Besstremyannaya, 2017a, Besstremyannaya and Golovan, 2019, 2021, 2022b, Besstremyannaya et al., 2022),
2. dynamic panel data models (Besstremyannaya, 2015, 2016, Besstremyannaya and Golovan, 2022b,c).

Unobserved heterogeneity

Unobserved heterogeneity is present when variables relevant for an agent's decision-making are unknown to the researcher (Browning and Carro, 2007).

Basic approaches for identification employ an instrumental variable model (Besstremyannaya, 2015b) or a fixed effect panel data model (Besstremyannaya, 2009a).

Examples of an *advanced approach* are the use of finite mixture (latent class) models:

- stochastic frontier analysis with finite mixtures (Besstremyannaya, 2011),
- linear finite mixture models (Besstremyannaya, 2015a, 2017b),
- binary choice models with finite mixtures (Besstremyannaya, 2017b),
- generalized finite mixture models (Besstremyannaya, 2015a, 2017b),
- finite mixture models for policy evaluation (Besstremyannaya, 2015a).

Overall, in view of empirical identification of heterogeneity in various economic settings, this dissertation develops and newly applies modern econometric techniques to econometrics in general and to several economics fields in particular. The indispensability of such an analysis may be supported by the failure of traditional models to explain numerous differences across economic agents[6] or the inability

[6] See Sen (2004), p.605 with examples about social differences in firm motivation which lead to heterogeneous productivity

of the conventional approaches to identify heterogeneous effects of policy reforms (Heckman, 2001).

3 Brief literature review

3.1 Finite mixture (latent class) models

The model is based on the assumption that an observation i comes from a mixture of a finite number of C unobserved classes (also often called components), and there are prior probabilities of belonging to each class $\pi_j > 0$, $j = 1, ..., C$, so that $\sum_{j=1}^{C} \pi_j = 1$.

Finite mixture models have been employed by mathematicians and statisticians since 1980s (Everitt and Hand, 1981, Clogg, 1981) but it was only in the late 1990s to early 2000s that these models became gradually introduced to microeconometrics (Compiani and Kitamura, 2016, McLachlan and Peel, 2000, Hagenaars and McCutcheon, 2002, Wedel and DeSarbo, 2002). The explanation of finite mixture models entered econometrics textbooks in the 2010s (Greene, 2012, Cameron and Trivedi, 2013), and only in 2020 a chapter dealing with the theory of mixture models appeared in *Handbook of Econometrics* (Schennach, 2020).

The earliest approach allowed for the simplest form of heterogeneity within finite mixture models: only the value of the constant term could differ across mixtures (Heckman and Singer, 1984). Applications of a more general approach with variation of all model parameters across classes originally focused on count data models and hurdle models (see review in Cameron and Trivedi (2013) and examples of the earliest works in health economics in Deb and Trivedi (1997) and Silva and Windmeijer (2001)). The analysis was soon extended to linear regression with continuous dependent variable (Deb and Holmes, 2000, Phillips, 2003), to stochastic frontier models (Greene, 2002, Tsionas, 2002), and generalized linear models (Greene, 2007).

The papers of this dissertation were the first to introduce a range of finite mixture models to health economics in general, as well as to the empirical analyses of costs of Japanese hospitals or healthcare expenditure of Japanese consumers, in particular (Besstremyannaya, 2017b, 2015a, 2011). The main features of a finite mixture model, which are employed in the papers of this dissertation, may be formulated as follows (Cameron and Trivedi, 2013, Greene, 2007).

The dependent variable y has the density

$$f(y_i|\pi, x, \theta) = \sum_{j=1}^{C} \pi_j f(y_i|x_i, \theta_j), \tag{1}$$

where x_i is the vector of explanatory variables for individual i and θ_j is the vector of unknown parameters associated with class j.

The method allows to model prior probabilities of class membership as functions of agent characteristics z_i. Under the assumption about the multinomial model for estimating prior class probabilities,[7]

growth; section 3.2 in Brock and Durlauf (2001) with examples from various economic fields, and section 4 in Kirman (2006) on heterogeneity in financial markets.

[7] The approach is commonly used in applications (Cameron and Trivedi, 2013, Greene, 2007, Bago d'Uva, 2005).

prior class probabilities may be expressed as

$$\pi_{ij} = \frac{exp(z_i'\gamma_j)}{\sum_{j=1}^{C} exp(z_i'\gamma_j)}, \text{ and a normalization is imposed through letting } \gamma_C = 0. \tag{2}$$

Bayes theorem is employed to estimate the posterior probability of i-th observation belonging to class j:

$$P(i \in j) = \pi_{ij} \cdot f(y_i|x_i,\theta_j) \Big/ \sum_{j=1}^{C} \pi_{ij} \cdot f(y_i|x_i,\theta_j) \tag{3}$$

Based on $\max\{P(i \in j)|j = 1,...,C\}$, the most probable class for each i is determined.

There is an implicit constraint through which the classes may be ordered, for instance in the case of two classes, $E(y_1|x) > E(y_2|x)$, so index 1 (class 1) stands for consumers with higher expenditure or firms with higher costs, and index 2 (class 2) denotes consumers with lower expenditure (or firms with lower costs).

Under the assumption of independent repeated measurements of y_{it} over time,[8] the joint density of y_{it} for the T repeated observations is the product of the marginal densities in each period:

$$f_j(y_i|\theta) = \prod_{t=1}^{T} f_j(y_{it}|\theta_{jt}), \tag{4}$$

where the marginal densities in periods with missing data are replaced by 1 (Wedel and DeSarbo, 2002, Greene, 2007).

Each observation is assumed to reside in the same class over the whole period of time, so

$$f(y_{it}|\pi,x_{it},\theta) = \sum_{j=1}^{C} \pi_{ij} \prod_{t=1}^{T} f(y_{it}|x_{it},\theta_j). \tag{5}$$

The estimate of the posterior joint probability of belonging to class j is:

$$P(i \in j) = \pi_{ij} \cdot \prod_{t=1}^{T} f(y_{it}|x_{it},\theta_j) \Big/ \sum_{j=1}^{C} \pi_{ij} \cdot \prod_{t=1}^{T} f(y_{it}|x_{it},\theta_j) \tag{6}$$

The choice of the number of classes C and the analysis of the goodness-of-fit in latent class models is commonly based on the comparison of residuals, information criteria (AIC, BIC), the value of the log-likelihood function, and chi-square tests (Andrews, 1988).

Different coefficients for the explanatory variables can be obtained for each class by using a regression analysis. Economic interpretation of the results is indispensable from the fact that the classes can be ordered with respect to the expected value of the dependent variable. However, the classes are probabilistic, so individuals can not be divided into groups with absolute certainty. Therefore, the analysis of posterior class probabilities (e.g. of belonging to the class with the highest expected value of the dependent variable) under this approach may be used only for tentative conclusions.

As regards the applicability of the finite mixture to the empirical analysis of a particular economic

[8]An assumption shared by the theoretical literature in the field (Wedel and DeSarbo, 2002, Bago d'Uva, 2005, Greene, 2007).

problem, much caution is required in the incorporation of this technique in actual economic settings: the justification of the use of the approach depends upon the plausible explanation for the existence of a discrete number of unobserved classes of economic agents.

3.2 Conditional quantile regression

3.2.1 Advantages of the approach

The conditional quantile regression model[9] allows to obtain independent estimates for the effect of covariates in each conditional quantile of the dependent variable. Therefore the approach may be regarded as superior to the conditional mean estimation, as it does not extrapolate the results of the mean regression to the tails of the distribution of the dependent variable. Different values of the estimated coefficients of an explanatory variable at regressions with different quantile indices are often interpreted as heterogeneity in the impact this covariate on the dependent variable. In other words, quantile regression analysis identifies heterogeneity through varying partial effects of the explanatory variable on the dependent variable. Examples of the applicability of quantile regression include a study of heterogeneous effect of macroeconomic shocks on firm costs, an evaluation of a varying impact of R&D investment on firm growth, an investigation of the differential effect of labor supply on worker wages.

The mapping function in conditional quantile regression is monotone (and in fact, strictly increasing), so the approach gives an ordered set of relationships. Therefore, the researcher may examine how the estimated coefficients for the explanatory variables change across the regressions with different (and ordered) values of the quantile index. As regards productivity analysis, high values of the quantile index (e.g. 0.8, 0.9) in the case of production function (output conditional on covariates) may be taken as an approximation of the production possibility frontier. When the conditional quantile regression is used for the estimation of cost function, low values of quantile index (e.g. 0.1, 0.2) may serve as an approximation for the best cost minimization trajectory. Accordingly, the quantile regression becomes applicable for efficiency analysis. For instance, the residual in the quantile regression for conditional output of the firm and $\tau = 0.8$ may be viewed as a measure of inefficiency.

3.2.2 Cross sectional model and pooled model

Denote $Q_\tau(y|x)$ as the conditional τ-th quantile of a continuous variable y under fixed values of explanatory variables x. The linear quantile regression regards the conditional τ-th quantile of a continuous variable y as a linear function of covariates x. The model originally appeared in Koenker and Bassett (1978).

The simplest longitudinal version of a quantile regression is a pooled model (Wooldridge, 2007). The Wooldridge (2007) correction of the variance matrix in such a pooled model enables to account for the serial correlation of errors within the clusters of observations. A formal proof of the asymptotic properties of such cluster-robust estimator may be found in Parente and Santos Silva (2016).

[9]This dissertation does not touch upon unconditional quantile regression models which are introduced by Firpo et al. (2009) as a powerful tool to study the impact of covariates on the *population* unconditional mean of the dependent variable. So the terms quantile regression and conditional quantile regression are used in this summary of dissertation interchangeably.

In this dissertation the pooled model with the Parente and Santos Silva (2016) cluster-robust standard errors is employed for estimating of the growth equation for Japanese manufacturing firms observed over a decade (Besstremyannaya et al., 2022). The use of the approach is justified by presence of the first lag of the dependent variable in the specification, which would inevitably cause endogeneity in the panel data model and limited applicability of the existing techniques for the analysis of panel data quantile regression models with endogeneity in case of very short panels.

3.2.3 Quantile-dependent fixed effects

A general form of a panel data quantile regression model is given in Koenker (2004):

$$y_{it} = x_{it}'\beta(U_{it}) + \alpha_i(U_{it}), \quad U_{it} \sim U[0,1], \tag{7}$$

$$\tau \mapsto x_{it}'\beta(\tau) + \alpha_i(\tau) \quad \text{is monotonically increasing,} \tag{8}$$

where $\tau \in (0,1)$, mapping (8) is the conditional quantile of the dependent variable y_{it}, x_{it} is a vector of covariates, i is the index for observation (as a longitudinal cluster), t is the index for time period, and $\alpha_i(\tau)$ are fixed effects, which vary across quantiles.

The conventional estimator is (eq.2.2. in Kato et al. (2012)):

$$(\{\hat{\alpha}_i\}, \hat{\beta}) = \underset{\{\alpha_i\},\beta}{\operatorname{argmin}} \frac{1}{nT} \sum_{i=1}^{n} \sum_{i=1}^{T} \rho_\tau(y_i - x_i'\beta - \alpha_i) \tag{9}$$

and its asymptotic theory requires long panels: n/T must be small.

The papers of this dissertation provide a review of the approaches that make it possible to estimate the model in case of short panels (Besstremyannaya and Golovan, 2021) and use one of the methods, i.e. the Galvao and Kato (2016) estimator, for measuring hospital productivity (Besstremyannaya and Golovan, 2022b). This becomes the first application of the general form of panel data conditional quantile regression model with fixed effects in health economics.

3.2.4 Quantile-independent fixed effects

The locational shift model assumes that fixed effects do not vary across quantiles. The model may be formulates as follows (Koenker, 2004).

$$y_{it} = x_{it}'\beta(U_{it}) + \alpha_i, \quad i = 1,\dots,n,\ t = 1,\dots,T, \tag{10}$$

where the function $\tau \mapsto x_{it}'\beta(\tau)$ is strictly increasing in τ, U_{it} is uniformly distributed on $[0,1]$ and does not depend on (x_{it}, α_i). Here x_{it} do not include the constant term. Individual effects α_i are considered as n additional unknown parameters.

In view of simplification of the computations, Canay (2011) proposed a simple estimator for the model with quantile-independent fixed effects.

$$y_{it} = x_{it}'\beta(U_{it}) + \beta_0(U_{it}) + \alpha_i, \quad i = 1,\dots,n,\ t = 1,\dots,T, \tag{11}$$

where the function $\tau \mapsto x'_{it}\beta(\tau) + \beta_0(\tau)$ is strictly increasing in τ, U_{it} is uniformly distributed on $[0,1]$ and does not depend on (x_{it}, α_i). Here the identification condition $E[\alpha_i] = 0$ is assumed. The first step of the approach consistently estimates the fixed effects with the help of any a $\sqrt{nT}$ consistent estimator (e.g. the within estimator). The second step clears the original dependent variable of the estimated fixed effects and applies the pooled version of the panel data quantile regression model to the new dependent variable.

The estimator is highly popular among practitioners as is shown in a meta-review in Besstremyannaya and Golovan (2019). This dissertation employs the estimator for the analysis of banking costs, thereby introducing the panel data conditional quantile regression approach to the banking literature (Besstremyannaya, 2017a). The panel employed in the paper is relatively long, i.e. the ratio of n/T is rather small and falls into the group of applications with the lowest value of this ratio (Besstremyannaya, 2017a, Besstremyannaya and Golovan, 2019). The dissertation also argues that the estimator may not be applicable for short panels (Besstremyannaya and Golovan, 2019) and provides recommendations for practitioners.

3.3 Dynamic panel data models

3.3.1 The justification for the use of the model

The dynamic panel data models may be regarded as a general form of the panel data fixed effect regression with the introduction of the dynamic structure of the data process. The dynamic structure (i.e. the lags of the dependent variable among covariates) is commonly justified by a habit-formation model due to behavioral features of consumers or firm managers.

Dynamic panel data models are often used in various economic fields. In macroeconomics, the methodology may be employed for the study of country growth (Bleaney et al., 2001, Laeven et al., 2015) or its current account (Wu, 2000). The use of dynamic panel data models in corporate finance and banking includes the analysis of firm size (Geroski et al., 1997, Oliveira and Fortunato, 2006), firm profit (Machin and Van Reenen, 1993) and bank profit (Knapp et al., 2006), firm leverage (González and González, 2012, Gaud et al., 2005), return on asset (by firms and by banks) and Tobin's Q (Kyereboah-Coleman, 2008, Goddard et al., 2004, Pérez-Calero et al., 2016). Financial applications deal with the study of prices (Gao et al., 2009, Santos, 2013). Research on labor, health and welfare uses the methodology of dynamic panel data models to study labor supply (Baltagi et al., 2005, Zhao et al., 2008), quality of healthcare (Mark et al., 2004), household wealth and individual health (Michaud and van Soest, 2008).

The papers of this dissertation use the methodology in health economics for adequate modeling of the quality of healthcare by hospitals (Besstremyannaya, 2016, 2015, Besstremyannaya and Golovan, 2022a,c) in order to measure heterogeneity in hospitals' response to price and quality regulation.

The simplest form of the dynamic panel data model is based on the assumption about the order 1 autoregressive process (Hamilton, 1994). The model may be extended to higher order lags, e.g. in case of the AR(2) process it becomes:

$$y_{it} - x'_{it}\beta - \mu = \alpha_1(y_{i,t-1} - x'_{i,t-1}\beta - \mu) + \alpha_2(y_{i,t-2} - x'_{i,t-2}\beta - \mu) + \nu_i + \varepsilon_{it}, \tag{12}$$

where i is the index for individual (as a longitudinal cluster), t indicates time, v_i are fixed effects and ε_{it} is i.i.d. random error.

Equation (12) and its extended analogue, which described in section 4.3.2 of this resume as equation (18), can be estimated using the generalized method of moments. Specifically, the Arellano and Bover (1995)/Blundell and Bond (1998) estimator, which provides for robust variance-covariance matrix (Windmeijer, 2005) has become one of the standard estimators in the literature. Overall, the instruments include lagged values of predetermined and endogenous variables (the first set of moment conditions) and differenced predetermined and endogenous variables (the second set of the moment conditions). The first set of moment conditions is based on the methodology of Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998), which takes the first difference of the right-hand side and left-hand side of equation (12). A second set of moment conditions is for the level equations (Blundell and Bond, 1998). The moment conditions are formulated separately for each year t.

3.3.2 On the quality of instruments

The validity of instruments is formally assessed with the help of the Arellano–Bond test. The Sargan test statistic may also be employed for evaluating validity of instruments, but it is applicable only under the homoskedasticity assumption.[10] But despite formal tests, the justification for the use of the instruments must be given on economic grounds. Specifically, the sets of moment conditions for dynamic panel data estimation employ lagged levels and lagged differences as instruments. However lags may turn out to be weak and invalid instruments (Bazzi and Clemens, 2013), particularly in case of distant lags (Murray, 2006) and in presence of overfitting of the endogenous variable in long panels (Roodman, 2009). So there may be problems with use of lags as instruments even though they pass the Arellano–Bond test.

Moreover, formal tests are only a necessary condition, and they can not fully establish the causal relationship in models, which use an instrumental variable approach (Angrist and Pischke, 2015, Bazzi and Clemens, 2013). So the justification of the assumption of the exclusion restriction of the instrument is provided in this dissertation on economic grounds (Besstremyannaya and Golovan, 2022a).

3.3.3 Mean reversion in dynamic panel data models

Dynamic panel data analysis is often used for evaluating the effect of a binary (or a continuous) variable, which may be regarded as a policy indicator (or policy intensity). For this purpose, the interaction term of the $y_{it-1} - x'_{it-1}\beta - \mu$[11] and of the variable for the participation in a reform, as well as a variable for the reform participation per se should be added to the right-hand of the equation (12). Using the estimated coefficients for the interaction term (and for the reform variable per se), it becomes possible to measure heterogeneity in the effect of the reform. Indeed, the effect depends on the values of y in the previous period(s) and can differ across observations with higher and lower values of y_{t-1} etc.

But such approach does not exclude the effect of the regression-towards-the mean (mean reversion), which is an integral part of the stationary process and hence, is inherent to dynamic panel data models (Dias and Marques, 2010, Gao et al., 2009, Knapp et al., 2006, Wu, 2000). Mean reversion may be

[10] So the approach is infeasible in the specifications with robust standard errors.
[11] And interactions of corresponding terms for higher order lags and the reform.

defined as the tendency of the stationary process to return to its mean value (Dias and Marques, 2010, Gao et al., 2009). The term "mean reversion" is believed to have first appeared in research by Galton which gave the start to applied econometrics – a set of papers on the inverse association between height of children and parents (Galton, 1886, Galton and Dickson, 1886). Examples of the processes with mean reversion may be found in various fields of economics: country productivity (Friedman, 1992), bank profitability (Knapp et al., 2006), prices of houses (Gao et al., 2009), blood pressure and cholesterol level of patients (Barnett et al., 2004).

This dissertation offers several approaches to overcome the problem of mean reversion in measuring policy response (Besstremyannaya, 2016, Besstremyannaya and Golovan, 2022a,c) and is novel in providing such approach for measuring the effect of time-varying policy reform as well as in giving an application in policy evaluation in health economics (Besstremyannaya and Golovan, 2022a).

3.4 Policy evaluation: average treatment effect, conditional average treatment effect and difference-in-difference estimation

The analysis of the effect of various economic policies has become a central part of modern econometrics (Angrist and Pischke, 2015, Heckman and Vytlacil, 2007, Heckman, 2001). The simplest method to account for the effect of the policy variable r on the dependent variable y is to include r in the list of regressors and look at the estimated coefficient for r. This way the "naive analyst" would obtain the difference in the values of the dependent variable across observations under the reform and not under the reform (Wooldridge, 2012, P.454). However, for non-randomized treatment assignment, the approach does not account for the potential differences in the characteristics of observations which have undergone the reform in comparison to other observations (Imbens, 2004).

Accordingly, a more careful approach would imply the construction of a control group of observations (also called counterfactuals) (Angrist and Pischke, 2009). It may be noted that "counterfactuals are required to forecast the effects of policies that have been tried in one environment but are proposed to be applied in new environments and to forecast the effects of new policies" (Heckman and Vytlacil, 2007, P.4782).

The analysis with counterfactuals appeared in the 1930s and saw a new wave of development in the 1960s–1980s (Heckman, 2001), owing to the introduction of modern approaches based on accurate definitions of the identification assumptions (see reviews in Heckman and Vytlacil (2007), Imbens (2004), Angrist (2004) and Imbens (2003) and examples in Angrist and Pischke (2015, 2009), Heckman (2001)).

An approach which has recently been gaining popularity in policy evaluation – the conditional average treatment effect estimation – is based on two major identifying assumptions. The first is the non-overlap of the treated and the controls (Angrist, 2004). The second is unconfoundedness – the conjecture that conditional on a set of covariates, participation in the reform does not depend on the outcome in each of the two states: participation and non-participation (Abadie and Imbens, 2016, Rosenbaum and Rubin, 1983). While a number of methods enable the construction a control group of observations for non-randomized trials, such as propensity score matching or inverse probability weights (Athey et al., 2018, Angrist and Pischke, 2015, Imbens, 2004, Hirano et al., 2003), there are many advantages associated with the use of nearest neighbor matching (Abadie and Imbens, 2002, 2006) as it does not depend on the smoothing parameters and allows an increase in the precision through raising the number of matches (Abadie and Imbens, 2011, Abadie et al., 2004).

As regards longitudinal data, the simplest method of evaluating the policy effect through the coefficient for the policy participation dummy variable suffers from contamination due to time trends and does not allow us to establish causality between the policy and the dependent variable (Angrist and Pischke, 2015, Wooldridge, 2012, Angrist and Pischke, 2009). A solution is the use of difference-in-difference estimations which evaluate the change in the dependent variable in the pre-reform and post-reform periods across the treated and the control observations. The similarity in the pre-reform trends across the treated and the controls is the necessary identifying assumption for difference-in-differences estimations.

Collapsing the data into pre-reform and post-reform period (Bertrand et al., 2004) offers a solution to the problem of the inconsistency of the standard errors in difference-in-difference estimations.

The approach is used in the papers of this dissertation in measuring the average treatment effect/conditional average treatment effect of the price reforms targeted at consumers and producers of healthcare (Besstremyannaya, 2013, 2017b). A statistical illustration of using policy evaluation techniques for assessing the impact of a reform in municipal finance is given in (Besstremyannaya, 2019a).

3.5 Non-parametric models

3.5.1 Parametric and non-parametric efficiency scores in frontier analysis

Frontier analysis stems from the seminal work of Farrell (1957), who offered definitions of the technical and price efficiency of a firm, and showed a method for constructing a linear convex hull surface to envelop observations. Essentially, the method regards, for instance, the technical efficiency of a firm as "the ratio of its mean production (conditional on its levels of factor inputs and firm effects) to the corresponding mean production if the firm utilized its levels of inputs most efficiently" (Battese and Coelli, 1992, P.154). Frontier analysis is often used to compute the efficiency scores of firms in order to employ them in the posterior estimations, such as policy evaluation.

A non-parametric method of frontier analysis is developed in Charnes et al. (1978) who proposed the term data envelopment analysis (DEA) – an approach aimed at finding a solution to a linear optimization problem. The original work of Charnes et al. (1978) deals with constant returns to scale and an input-oriented model but the framework was later extended to variable returns to scale, non-increasing returns to scale, output-oriented and other DEA models (Banker et al., 1984, Färe et al., 1985, Seiford, 1996). The DEA efficiency score is often used as a dependent variable in regression analysis which studies heterogeneous effect of various characteristics of the economic agent on the agent's efficiency. Moreover, data envelopment analysis per se is regarded as a part of both operational research theory and econometric analysis (Tone, 2017, Ray, 2004), and papers in *Journal of Econometrics* deal with the theory and applications of the methodology (Simar and Wilson, 2007, Cazals et al., 2002, Charnes et al., 1990, Seiford and Thrall, 1990). A theoretical issue raised in the dissertation with respect to data envelopment analysis is the need for the bias-correction of scores obtained within cost minimization DEA.

An alternative parametric method, stochastic frontier analysis (SFA) employs a composite error term in either the production or cost function equation: it is the sum of statistical noise and the inefficiency component (Aigner et al., 1977, Battese and Corra, 1977, Meeusen and van Den Broeck, 1977). The advantages of DEA are as follows: the ability to deal with multi-output production functions, no assumptions about the functional form of production or cost function and non-vulnerability to the problem of multicollinearity. However, DEA is sensitive to outliers, and other disadvantages of the approach are the inability to incorporate measurement error, and the fact that efficiency scores are exactly unity for the firms on the constructed frontier.

The parametric method, SFA, is able to account for measurement errors and outliers through statistical noise, and does not require that the efficiency score equals unity. The weak points of SFA are rigidity in the assumptions about the form of the production (or cost) function and the distributions of the error term. The method does not explicitly account for multi-output functions.

Accordingly, DEA and SFA are often used as complementary methods for estimating efficiency scores (Kooreman, 1994, Seiford and Thrall, 1990). In this dissertation, DEA and SFA scores are employed for measuring efficiency of Japanese hospitals in order to evaluate the impact of a hospital financing reform on efficiency (Besstremyannaya, 2013).

3.5.2 Kernel regressions in contrast to OLS models

While a parametric model inevitably imposes restrictions on the economic process, kernel density estimators do not deal with functional form assumptions and therefore, may be considered a convenient tool for applied non-parametric analysis with a large sample size and a limited number of regressors (Härdle and Linton, 1994). Examples of early works are Wang and Van Ryzin (1981) and Van Ryzin and Wang (1978), while recent extensions include papers by Parmeter and Racine (2019), Racine (2019), Hayfield and Racine (2011, 2008), Li and Racine (2008), Hsiao et al. (2007), Racine and Li (2004), Li and Racine (2003).

In this dissertation kernel regression is used as a complementary technique to OLS regression in order to estimate the impact of the private health insurers on the quality/effectiveness of regional healthcare systems in Russia (Besstremyannaya, 2017b).

4 Methodology

4.1 Finite mixture models

4.1.1 Stochastic frontier model with latent classes

Besstremyannaya (2011) is the first paper in the health economics literature that captures unobserved heterogeneity in hospital costs through a stochastic frontier model with latent classes. The purpose of the analysis is to establish a link between unobserved managerial practices and the cost efficiency of hospital as a firm. The approach is based on the premise of the general availability of state-of-art technology by firms in each industry. Therefore, inefficiencies may be due to ineffective managerial practices (Bloom et al., 2016) as management may be regarded as an inseparable part of production (Bloom and Van Reenen, 2007).

Using an example of Japanese acute-care local public hospitals in 1999–2007, the paper hypothesizes that differences in managerial practices that are established for manufacturing firms and for hospitals in different countries (Bloom and van Reenen, 2010, Bloom and Van Reenen, 2007) divide hospitals into an unobserved and finite number of classes with respect to managerial practices. These practices affect hospital costs, and therefore, cost minimization trajectories (and cost efficiency scores) must be measured separately for each class. The analysis applies the Greene (2002) and Tsionas (2002) stochastic frontier model with latent classes to the estimation of a hospital cost function as follows.

$$\begin{aligned}\ln\frac{c_{it}}{p_{Kit}}|j = \sum_{m=1}^{M}\beta_{mj}\ln y_{mit} + \sum_{k=1}^{K-1}\beta_{kj}\ln\frac{p_{kit}}{p_{Kit}} + 0.5\sum_{s=1}^{K-1}\sum_{o=1}^{K-1}\beta_{osj}\ln\frac{p_{oit}}{p_{Kit}}\ln\frac{p_{sit}}{p_{Kit}} \\ + 0.5\sum_{m=1}^{M}\sum_{n=1}^{M}\beta_{mnj}\ln y_{mit}\ln y_{nit} + \sum_{k=1}^{K-1}\sum_{m=1}^{M}\beta_{kmj}\ln\frac{p_{kit}}{p_{Kit}}\ln y_{mit} + \sum_{j=1}^{J}\beta_{lj}z_{lit} + v_{itj} + u_{itj}\end{aligned} \tag{13}$$

$$u_{itj} = exp\{-\eta_j(t-T)\}\cdot U_{ij} \geq 0,\ (\eta_j < 0 \text{ is increasing inefficiency}) \tag{14}$$

$$U_{ij} \sim N^+(0,\sigma_{uj}^2) \tag{15}$$

$$v_{itj} \sim N(0,\sigma_{vj}^2) \tag{16}$$

where c_{it} is total costs of i-th hospital in period t, $y_{it} = (y_{1it},...,y_{Mit})$ is a vector of outputs that account for the inpatient and outpatient activity of a hospital (the number of outpatient visits and the mean number of admissions and discharges), $z_{it} = (z_{1it},...,z_{Lit})$ is a vector of hospital characteristics, $p_{it} = (p_{1it},...,p_{Kit})$ is a vector of input prices, p_{Kit} is a numeraire price, j is index for a latent class, η is the Battese and Coelli (1992) time decaying parameter.

Managerial performance indicators are used in the posterior analysis of latent class membership.

4.1.2 Generalized finite mixture models

Besstremyannaya (2017b, 2015a) are the first papers to study the demand for healthcare that apply generalized finite mixture models to account for unobserved differences of individuals. The analysis is targeted at investigating the interrelation between the behavioral characteristics of individuals and their demand for healthcare services and drugs. The demand is approximated by healthcare expenditure, which is commonly analyzed as the logged dependent variable (McCullagh and Nelder, 2019). So the econometric novelty of the two papers is in applying the Greene (2007) approach with a combination of generalized linear models for panel data with a logged dependent variable and finite mixture models.

The novelty of the analysis in economic terms is the study of price elasticity and income equity by consumers with higher and lower healthcare expenditure. The major assumption in the analysis is that the behavioral and health characteristics of consumers separate them into unobserved classes, with different price and income elasticity of demand for healthcare in classes of “high users” and “low users” of healthcare.

The generalized linear model is

$$f(E(y_{it}|x_{it},j)) = x_{it}'\delta_j, \text{and } (y|x_{it},j) \sim g(y_{it},x_{it},\theta_j), \tag{17}$$

where f is a link function, g is a family of distribution, δ_j are coefficients, θ_j are parameters for j-th class, and y_{it} is healthcare expenditure by individual i in period t.

Goodness of fit analysis and the choice of the number of latent classes is conducted in the two papers with the help of residual analysis in each class (raw bias, absolute prediction error, squared error), information criteria, Andrews (1988) chi-squared test and cross-validation (50 replications with a randomly chosen 80% of observations as a training sample and the remaining 20% as a holdout sample at each replication).

The model in Besstremyannaya (2015a) is applied to panel data for adult Japanese consumers in 2008–2010 and young adult female Japanese consumers in 2002–2010, while the analysis in Besstremyannaya (2017b) deals with 2009–2014 panel data for Japanese adults.

4.1.3 Binary choice models with latent classes

Besstremyannaya (2017b) accounts for unobserved consumer heterogeneity in measuring income equity in access to healthcare. The analysis employs binary choice models with latent classes, which are an extension of the Deb and Trivedi (2002) model and an application of the Bago d'Uva (2005) approach.

The novelty of the paper is twofold. Firstly, it uses panel data binary choice finite mixture models and separately examines the use of outpatient and inpatient healthcare. Secondly, it measures income equity within the Japanese social health insurance system under unobserved heterogeneity of consumers.

4.2 Conditional quantile regression

4.2.1 Empirical analysis

The section on finite mixture (latent class) models demonstrated how the approach can help analyze unobserved heterogeneity of agents, for instance, in the study of the interrelation between managerial practices and costs of firms. However, the methodology has its limitations: the classes are probabilistic and the number of classes is discrete. The conditional quantile regression offers alternative means to study differences across firms that vary in their productivity, costs or growth. The papers in this dissertation apply the methodology to the analysis of banking costs (Besstremyannaya, 2017a), hospital production (Besstremyannaya and Golovan, 2022b) and the growth of innovative manufacturing firms (Besstremyannaya et al., 2022). The approach may be also employed to analyze heterogeneity across consumers and an example is the study of the heterogeneous effect of endogenous labor supply on wages in the extensions section of the resume and the dissertation (Besstremyannaya and Golovan, 2022).

Besstremyannaya (2017a) is the first paper to employ a conditional quantile regression for the analysis of longitudinal data on banking costs. The paper assumes that differences in managerial and business practices at banks[12] have consequences for the bank's ability to minimize costs and sustain exogenous shocks. Accordingly, the purpose of the paper is to reveal the heterogeneous effect of macroeconomic shocks on costs of low-cost and high-cost banks and to discover differences in their economies of scale and scope. The analysis uses a nationwide sample of over 100 Japanese banks in 2001–2013 and employs the Canay (2011) panel data model with quantile-independent fixed effects.

[12] As revealed in numerous papers, see reviews in (Hughes and Mester, 2013, Caprio and Honokan, 2014).

Covariates in the multi-product cost function include banking outputs which account for the attitude to risk and specific features of the Japanese banking system. Control variables are employed at the level of the bank and the level of the prefecture. Second-stage analysis is applied to the residual in conditional quantile regressions with low values of τ, and the residual is treated as a measure of cost inefficiency. Cost inefficiencies are explained by the list of variables related to capital structure, bank risk, and profitability. The analysis uses generalized method of moments as it enables us to account for endogeneity.

Besstremyannaya and Golovan (2022b) follow the approach of Besstremyannaya (2017a) using a conditional quantile regression for the analysis of the heterogeneity of firm production/costs. The paper is the first research in health economics to employ a conditional quantile regression to the analysis of panel data on hospital production. The study deals with public hospitals where soft budget constraints make the analysis of production more appropriate than the analysis of costs (Biørn et al., 2010). Accordingly, the paper estimates the extended version of the multi-output production function of hospitals which was employed in Besstremyannaya (2013). Specifically, hospital outputs are measures of inpatient and outpatient activity, as well as a proxy for research activity by hospital personnel. The paper uses the model with quantile-dependent fixed effects. As it deals with short panels, it employs the smoothing techniques by Galvao and Kato (2016) and Dhaene and Jochmans (2015).

The purpose of the paper is to disentangle differences in the productivity of capital and labor, and to evaluate the optimality of the labor/capital mix at high-output and low-output hospitals. The paper assumes that ineffective management leads to differences in the elasticity of hospital output with respect to production factors. It focuses on local acute-care public hospitals in Japan in 1999–2019 and evaluates differences in the productivity of hospital inputs (labor specialties, capital and medicines), and in the partial effects of hospital variables on hospital output. In the second-stage analysis the paper measures the production efficiency (as a residual in regression with high τ) and establishes an association between efficiency and a range of regional and municipal variables. Finally, the paper carries out a counterfactual policy analysis: an evaluation of potential cost savings in case of changeover to optimal values of hospital inputs.

Besstremyannaya et al. (2022) focus on the differences in R&D management which may be linked with different relationships between R&D intensity (R&D-to-sales ratio) and the growth of innovative firms. The paper applies a conditional quantile regression to study the heterogeneity of firm growth and uses nationwide samples of high-tech manufacturing firms in Japan in 2009–2020. The analysis is the first study on the heterogeneous effect of the R&D intensity of the growth of Japanese firms. It uses a longitudinal version of the conditional quantile regression model to estimate the augmented Gibrat law equation for each of four innovative industries: chemicals and allied products; electronic and other electrical equipment; industrial and commercial machinery and computer equipment; and transportation equipment. The paper follows the approach in Besstremyannaya and Golovan (2022a) to measure the efficiency of a firm as a residual in a regression with high τ.

4.2.2 Theoretical issues

Theoretical issues raised in the dissertation deal with a conditional quantile regression for panel data. Besstremyannaya and Golovan (2019) review the panel data model with quantile-independent fixed effects and Besstremyannaya and Golovan (2021) discuss the smoothing technique in the estimation of short panels in cases of quantile-dependent and quantile-independent fixed effects. While these two papers in the main part of the dissertation consider conditional quantile regressions under exogeneity, as extensions of the research, Besstremyannaya and Golovan (2022) touch upon the estimator with cluster-robust standard errors in cases of endogeneity.

Besstremyannaya and Golovan (2019) study the applicability of the computationally simple estimator by Canay (2011) with quantile-independent fixed effects. Given two errors in the estimator (i.e. firstly, the insufficiency of the condition $n/T^s \to 0$, where n is the number of longitudinal clusters of observations, T is the length of panel, and $s \in (1, \infty)$ for the asymptotic unbiasedness or existence of the estimator of the vector of the coefficients, while the condition $n/T \to \infty$ is more appropriate, and secondly, the fact that the standard error for the constant term is inestimable), the paper provides recommendations to practitioners on the use of the estimator. It argues that the estimator cannot be employed for large values of n/T. The applicability of the estimator in the case of a small n/T and regressors correlated across time periods requires the use of a bootstrap to solve issues with the estimation of the standard errors for the vector of coefficients. If n/T is small and regressors are independent across time periods, a bootstrap is required only for the estimation of the standard error of the intercept.

The 13-page appendix to Besstremyannaya and Golovan (2019) provides a meta-review of all empirical papers listed on the publisher's webpage of *Econometrics Journal* (as of the end of December 2018) that employ the Canay (2011) estimator. The number of papers is 81 and the detailed summary table outlines the economic field, the type of observation (i.e. firm, country, industry, household, individual, pairs of countries, employee-employer pairs etc.), the values of n and T, and whether the paper employed a bootstrap for the estimation of the standard errors. The applied analyses often used short panels ($T < 10$ in almost half of the papers) and the value of n/T is large (over 10 in over 70% of papers). Many papers do not estimate standard errors with a bootstrap nor touch upon the issue of the independence of regressors over time.

Besstremyannaya and Golovan (2021) review the approaches for estimating longitudinal models for conditional quantile regression. The paper highlights the fact that a method of smoothed quantile regression may be viewed as a remedy for reducing the asymptotic bias of the estimator in short panels. As regards the estimation of a quantile-dependent fixed effect model with short panels, it was originally implemented through imposing various restrictions: e.g., assumptions about the distribution of the dependent variable (Machado and Santos Silva, 2019, Li et al., 2003) or about the functional form of the fixed effects (Harding and Lamarche, 2016). However, the smoothing technique proposed by Galvao and Kato (2016) offers a solution for estimating the general form of the fixed effect quantile regression model. The Koenker (2004) quantile regression objective function is modified through smoothing as follows: $\min_{\{\alpha_i\},\beta} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - x_{it}'\beta - \alpha_i)(\tau - G(y_{it} - x_{it}'\beta - \alpha_i)/h)$, where $G(v) = \int_u^\infty K(v)dv$ is a smoothed analog of the step function $I(u \geq 0)$, $K(v)$ is a kernel function, h is the bandwidth, i indicates individual and t denotes time.

The smoothing technique of Galvao and Kato (2016) enables us to obtain the bias of the estimator, and two methods are suggested for the reduction of the bias: deducting the asymptotic expression of the bias or employing the Dhaene and Jochmans (2015) jackknife split panel correction of the bias.[13] The

[13]As regards balanced panels, the Dhaene and Jochmans (2015) procedure splits the panel into two: $i \in \{1, \ldots, n\}$ in each

split panel estimator under the Dhaene and Jochmans (2015) approach has the asymptotic variance equal to the variance of the original $\hat{\beta}$ estimator, and therefore enables a reliable inference for short panels.

In the case of quantile-independent fixed effects and in view of the asymptotic bias of the Canay (2011) estimator shown in Besstremyannaya and Golovan (2019), Chen and Huo (2021) constructed a new estimator for quantile-independent fixed effects specified by (11). Their estimator employs the normalization condition $E[\beta_0(U_{it})] = 0$ and exploits the first-step of the Canay (2011) procedure. However, the second step is modified with the help of the smoothing technique of Galvao and Kato (2016) which allows a reduction of the asymptotic bias of the estimator in short panels.

4.3 Dynamic panel data

4.3.1 Empirical analysis

The papers in this dissertation employ dynamic panel data models in order to correctly describe the dependent variable and to evaluate the heterogeneous effect of price and incentive regulation (with an application to inpatient healthcare in Japan and the US). The assumption about the heterogeneity of the effect stems in each case from the specific design of each incentive scheme. Proxies for the unobserved quality of hospital services are early readmission rate (readmission within 42 days after discharge in case of Japan), clinical indicators of the process of care, and patient assessment of the quality of care and outcomes of care (at US Medicare hospitals). Additionally, the average length of inpatient stay is used as a proxy for cost inefficiency in Japanese hospitals. (Shorter stays of patients are associated with lower costs and hence may be interpreted as a reflection of better managerial efforts and lower cost inefficiency.) The main reason for the dynamic nature of the process in cases of quality of care or length of stay is habit-formation on the part of hospital management and personnel as well adherence to the hospital-specific treatment patterns.

Besstremyannaya (2016) deals with a notable example of a financial reform in the regulated industries: the changeover from retrospective remuneration of incurred costs (fee-for-service) to prospective payment system (PPS) with a fixed payment.[14] Variants of PPS were introduced into healthcare systems of numerous countries as means to curb costs and raise production (Street et al., 2011). Japanese PPS follows the common rule of using an extension of the mechanism in the form of a two-part tariff (a fixed price component and a component dependent on the actually incurred costs, see (Laffont and Tirole, 1993). A special feature of the Japanese PPS is the per diem step-down schedule, with a daily fixed payment inversely dependent on the length of stay.

The theoretical model in Besstremyannaya (2016) shows that such a system provides disincentives for hospitals with the lowest pre-reform length of stay: their length of stay is likely to increase. The length of stay at hospitals with median and high pre-reform values should go down. However, the fall in

panel, while the time index is $t \in \{1, \dots, T/2\}$ in the first panel and $t \in \{T/2+1, \dots, T\}$ in the second panel. The split panel estimator is calculated as $\hat{\beta}_{1/2}(\tau) = 2\hat{\beta}(\tau) - (\hat{\beta}_1(\tau) + \hat{\beta}_2(\tau))/2$, where $\hat{\beta}(\tau)$, $\hat{\beta}_1(\tau)$, $\hat{\beta}_2(\tau)$ are respectively, estimators for the full panel, the first part of the panel and the second part of the panel.

[14] The reform is motivated by the Shleifer (1985) yardstick competition approach and the Laffont and Tirole (1993) model of regulation in public procurement. In its application to inpatient healthcare, the mechanism is based on diagnosis related groups (DRGs), carefully developed as "a system of describing hospital production" (Fetter and Freeman, 1986). Started on a trial basis in selected states of the US and then applied to all Medicare hospitals, this innovative system reimburses a fixed amount for treating a patient with a given DRG. Essentially, this becomes a fixed price contract on the quantity of services.

the quality of services at these hospitals may be an adverse effect of the reform. To sum up, the Japanese version of PPS is likely to have a heterogeneous impact on cost efficiency and the quality of hospital care. For each group of the diagnoses a proxy for cost inefficiency – length of stay – is likely to increase at hospitals with the lowest length of stay and to fall at hospitals with a median and high length of stay. The quality of care may go down at hospitals with a median and high length of stay.

The empirical analysis in the paper employs order 1 dynamic panel data models to estimate the equations for the length of stay and the early readmission rate at the level of major diagnostic categories[15] in 1068 Japanese hospitals in 2006–2012. The change in the fitted value of the dependent variable before and after the reform is then studied using the pre-reform quartile of the dependent variable. The approach does not enable us to disentangle the effect of the mean reversion in measuring the impact of the reform. To alleviate the potential impact of mean reversion, the analysis is supplemented by the estimation of annual cross-section versions of the dynamic panel data model, with the first lag of the dependent variable included in the list of covariates. Another attempt to mitigate the problem of mean reversion is the computation of the mean value of the dependent variable in several post-reform years. This value is then compared to the value before the reform.

Besstremyannaya and Golovan (2022c) and Besstremyannaya and Golovan (2022a) consider the incentive contract on quality.[16] It is the so-called pay-for-performance mechanism – an innovative method of remuneration, which originally emerged in corporate finance and managerial economics, and has since been widely used in the public sector (civil service, education, social work, and healthcare). In order to quantify the unobserved quality of work, the incentive scheme computes the performance level using imprecisely measured proxies for various dimensions of quality (analog of key performance indicators at firms). Next, the regulator imposes an (intertemporal) incentive contract which associates remuneration with performance. In this way, agents with higher performance in the current period receive a higher payment for their services in future periods than agents with lower performance. In US Medicare, the reform was implemented in 2013 on the basis of a reward function that linearly links the total performance score (an aggregate measure of hospital quality) to remuneration for acute inpatient care. So a PPS contract on quantity is supplemented with a contract on quality.

The total performance score is the weighted sum of scores for measures in several domains: the timely implementation of recommended medical interventions (the clinical process of care), the quality of healthcare as perceived by patients (the patient experience of care), the survival rates for AMI, heart failure, and pneumonia patients, and other proxies for outcome of care, healthcare-associated infections and other measures of safety of care, and spending per patient as a measure of the efficiency of care.

The available data allows us to estimate order 2 dynamic panel data models for the total performance score or its dimensions. As regards the heterogeneity of the incentive, Besstremyannaya and Golovan (2022c) focus on the payment schedule related to the total performance score. Owing to the linear character of the incentive, the paper hypothesizes that the effect of the reform (defined as increase of quality) is likely to be larger for hospitals with a higher pre-reform value of total performance score. The empirical part of the paper deals with a sample of 3,000 Medicare hospitals in 2011–2019. (It should be noted that while the components of the aggregate measure are computed since 2004, the aggregate

[15]Constructed on the basis of international classification of diseases

[16]Both papers elaborate on the approach of Besstremyannaya (2015).

measure is available only since 2011.) The paper explicitly estimates the effect of mean reversion by using the fact that the intensity of the reform (the share of hospital's funds at risk) varies across years. So the paper computes the unconditional mean in the autoregressive process as a function of the hospital's characteristics and the reform intensity. A comparison of the fitted values of the unconditional mean under different values of the reform intensity enables the identification of the policy effect cleared of mean reversion.

The validity of the instruments in the system GMM estimations is discussed on economic grounds. According to the interviews with hospital managers, hospital administrators and personnel take prompt action upon learning the total performance score in year t (the decision-making is conducted not on the annual but on the quarterly or even monthly basis). So adjustment in the value of the aggregate quality measure occurs in period $t+1$ and is not delayed until a more remote future. This way the instruments, i.e. Δy_{t-1} as an instrument for y_t in equation (18), are unlikely to affect the dependent variable though other channels than the endogenous variable and potentially, hospital control variables.

Besstremyannaya and Golovan (2022a) focus on each dimension of the aggregate quality measure in the US pay-for-performance mechanism at 3,000 Medicare hospitals in 2004–2017. Developing a theoretical model about the effect of quality incentives on altruistic and motivated providers,[17] the paper predicts that the impact of the reform may be heterogeneous across different observed measures of hospital quality and across hospitals with varying pre-reform values of these measures. The model forecasts a crowding out of the most altruistic types. Next, the paper assumes that altruism is heterogeneous across hospitals and the values of altruism in each hospital are higher for quality measures which are strongly associated with the patient's benefit. The analysis employs dynamic panel data estimations where the reform is treated as a binary variable (i.e. the value of the reform intensity is not taken into account). So the empirical approach excludes pre-reform and post-reform ''regression-to-the-mean" effects by modeling the pre-reform and post-reform long-term means as a function of the hospital characteristics. The effect of the reform is then measured as the sum of the estimated coefficient for the reform dummy (i.e. the effect at the unity value of the reform variable) and of the estimated coefficients for the interaction terms of the reform dummy and other variables (at mean values or at decile groups of the corresponding variables and the unity value of the reform variable).

4.3.2 Theoretical issues

To assess the effect of reform cleared of mean reversion, Besstremyannaya and Golovan (2022c) compute the long-term mean μ in the autoregressive process as a function of hospital characteristics and the intensity parameter of the reform α_t. If the reform intensity varies over time, it becomes possible to estimate the impact of the reform on the long-term mean. Two variants of the estimation are proposed in the paper: 1) by plugging in the values of the intensity at period t and $t+1$ and computing the difference between $\mu(\alpha_{t+1})$ and $\mu(\alpha_t)$ and 2) by computing the value of $\mu(\alpha_t) - \mu(0)$.[18] To the best of our knowledge, the only related approach may be found in Knapp et al. (2006) who propose a below

[17] The extended working paper version of the article provides meta-review of the experimental and empirical literature on the existence of the altruism on healthcare markets (Besstremyannaya and Golovan, 2019).

[18] Similarly, by estimating the persistence parameter λ as a function of the time-varying intensity of reform, it becomes possible to evaluate the impact of the reform on λ.

method for evaluating the policy effect of bank mergers in dynamic panels: the actual value of return on equity (ROE) in merged banks is compared with the fitted value of ROE, measured as the unconditional mean of the AR(1) process for the whole banking industry (i.e. the counterfactual value of ROE in the absence of the merger, equivalent to a zero value of the policy intensity parameter in our application below).

In an application to Medicare's incentive contract, y_{it} is the total performance score of hospital i in year t. The paper uses the second order dynamic panel:

$$y_{it} = \phi_0 + \phi_1 y_{it-1} + \phi_2 y_{it-2} + \phi_3 \alpha_t s_{it} + \phi_4 \alpha_t s_{it} y_{it-1} \\ + \phi_5 \alpha_t s_{it} y_{it-2} + \delta_0 s_{it} + z'_{it}\delta_1 + \alpha_t s_{it} \cdot z'_{it}\delta_2 + d'_t\delta_3 + u_i + \varepsilon_{it}, \quad (18)$$

where α_t is the size of the quality incentive (a time-varying parameter of the reform intensity), s_{it} is the share of Medicare discharges, z_{it} are time-varying hospital characteristics, u_i are fixed effects and ε_{iy} is i.i.d. random error.

For a fixed value of α, the paper takes the unconditional expected values of both sides of (18) and denotes $\mu(\alpha) = E(y_{it})$:

$$\mu(\alpha) = \frac{\phi_0 + \phi_3 \alpha E(s_{it}) + \delta_0 E(s_{it}) + E(z_{it})'\delta_1 + \alpha E(s_{it}z_{it})'\delta_2 + \phi_4 \alpha \operatorname{cov}(s_{it}, y_{it-1}) + \phi_5 \alpha \operatorname{cov}(s_{it}, y_{it-2})}{1 - \phi_1 - \phi_2 - \phi_4 \alpha E(s_{it}) - \phi_5 \alpha E(s_{it})}. \quad (19)$$

Since α differs across t, so the paper uses sample means across the hospitals for fixed t to obtain estimates of expectations. The estimate of $\mu(\alpha)$ is constructed by replacing the expected values and covariances by corresponding sample means and sample covariances:

$$\mu(\alpha) = \frac{\phi_0 + \phi_3 \alpha \bar{s} + \delta_0 \bar{s} + \bar{z}'\delta_1 + \alpha \overline{sz}'\delta_2 + \phi_4 \alpha \widehat{\operatorname{cov}}(s, L(y)) + \phi_5 \alpha \widehat{\operatorname{cov}}(s, L^2(y))}{1 - \phi_1 - \phi_2 - \phi_4 \alpha \bar{s} - \phi_5 \alpha \bar{s}}.$$

For a second order autoregressive process, the difference between the conditional expected value of y_{it} and the long-term mean decays exponentially at the rate equal to the reciprocal value of the smallest root of the characteristic equation for the process (Hamilton, 1994, Section 2.3):

$$1 - (\phi_1 + \phi_4 \alpha_t s_{it})\lambda - (\phi_2 + \phi_5 \alpha_t s_{it})\lambda^2 = 0.$$

So, for a fixed value of α the paper takes the expectations

$$1 - (\phi_1 + \phi_4 \alpha E(s_{it}))\lambda - (\phi_2 + \phi_5 \alpha E(s_{it}))\lambda^2 = 0.$$

Then the expected values are replaced by sample means. Solving this quadratic equation, we obtain:

$$\lambda(\alpha) = \frac{\phi_1 + \phi_4 \alpha \bar{s} + \sqrt{(\phi_1 + \phi_4 \alpha \bar{s})^2 + 4(\phi_2 + \phi_5 \alpha \bar{s})}}{2},$$

where $\bar{s}$ is the mean value of s for a given year.

The policy parameter α in the US Medicare reform is zero before 2013, increases in 2013–2017 and remains unchanged in 2017–2019. To examine the hypothesis about the positive effect of the reform, we can look at the difference between $\mu(\alpha_t)$ and $\mu(\alpha_{t-1})$:

$$\begin{aligned}\mu(\alpha_t)-\mu(\alpha_{t-1}) = & \frac{\phi_0+\phi_3\alpha_t\bar{s}+\delta_0\bar{s}+\bar{z}'\delta_1+\alpha_t\overline{sz}'\delta_2+\phi_4\alpha_t\widehat{\text{cov}}(s,L(y))+\phi_5\alpha_t\widehat{\text{cov}}(s,L^2(y))}{1-\phi_1-\phi_2-\phi_4\alpha_t\bar{s}-\phi_5\alpha_t\bar{s}} \\ & -\frac{\phi_0+\phi_3\alpha_{t-1}\bar{s}+\delta_0\bar{s}+\bar{z}'\delta_1+\alpha_{t-1}\overline{sz}'\delta_2+\phi_4\alpha_{t-1}\widehat{\text{cov}}(s,L(y))+\phi_5\alpha_{t-1}\widehat{\text{cov}}(s,L^2(y))}{1-\phi_1-\phi_2-\phi_4\alpha_{t-1}\bar{s}-\phi_5\alpha_{t-1}\bar{s}}.\end{aligned}$$

The null hypothesis is: $H_0\colon\ \mu(\alpha_t)-\mu(\alpha_{t-1})=0$, and it is tested against the positive alternative.

Equivalently, it is possible compute the difference between $\mu(\alpha_t)$ and $\mu(0)$ (index t is omitted in the expression below):

$$\begin{aligned}\mu(\alpha)-\mu(0) = & \frac{\phi_0+\phi_3\alpha\bar{s}+\delta_0\bar{s}+\bar{z}'\delta_1+\alpha\overline{sz}'\delta_2+\phi_4\alpha\widehat{\text{cov}}(s,L(y))+\phi_5\alpha\widehat{\text{cov}}(s,L^2(y))}{1-\phi_1-\phi_2-\phi_4\alpha\bar{s}-\phi_5\alpha\bar{s}} \\ & -\frac{\phi_0+\bar{z}'\delta_1}{1-\phi_1-\phi_2}.\end{aligned}$$

The null hypothesis is: $H_0\colon\ \mu(\alpha_t)-\mu(0)=0$, and it is tested against the positive alternative.

4.4 Parametric and non-parametric analysis

4.4.1 Empirical analysis with parametric and non-parametric efficiency scores

Besstremyannaya (2013) computes parametric and non-parametric efficiency scores of Japanese acute-care local public hospitals in order to employ them for the second-stage analysis and evaluate heterogeneity in the response of hospital efficiency to a financing reform (an introduction of inpatient prospective payment system). The parametric technical efficiency scores are computed with the help of stochastic frontier analysis models, applied to the multi-output production function of hospitals. Output-oriented data envelopment analysis (DEA) model is used to compute a non-parametric analog of the SFA technical efficiency. This follows the standard approach in the literature which regards SFA and DEA as complementary techniques (Nakayama, 2003, Jacobs, 2001, Kooreman, 1994). Along with the output-oriented efficiency, the paper also computes cost efficiency through Tone (2002) cost-minimization DEA and suggests an approach for bias-correction of these cost-efficiency scores. The impact of the hospital financing reform on non-parametric and parametric efficiency scores is then studied through descriptive analysis (Besstremyannaya, 2013, Table 5) and linear regression in difference-in-difference estimations (described in the Policy evaluation section of this summary).

Besstremyannaya and Simm (2019) employ cost minimization DEA model to compute efficiency scores of Japanese banks and conduct a second-stage analysis in order to find heterogeneity with respect to bank characteristics and other control variables. The estimation includes the model with environmental variables:[19] a two-stage approach of Simar and Wilson (1998) and a one-stage approach of Simar and Wilson (2007). For robustness, the analysis considers the banking production under the asset approach and under intermediation approach, and uses two measures of cost efficiency: developed in Färe et al.

[19] Bank variables and prefectural variables not considered as inputs in the model, e.g. binary variable for bank charter, index of product diversity, rate of growth of regional GDP, share of monetary aggregate in GDP etc.

(1985) and in Tone (2002). Although the coefficients for environmental variables are not reported in the paper, the REPEC working paper versions present plots of cost-efficiency scores with an indication of the bank charter (Besstremyannaya et al., 2017, P.20) and (Besstremyannaya and Simm, 2015, P.19).

4.4.2 Parametric and kernel regressions

Besstremyannaya (2015b) applies parametric and kernel regressions to study the impact of private health insurers on quality-related outcomes of mandatory health insurance in Russian regions. The analysis uses the simplest approach for making judgment about heterogeneous impact of a covariate: it examines whether the estimated coefficient for this variable is significant. The proxies for the quality of health system are infant mortality, maternal mortality and mortality under five years of age. The parametric analysis uses OLS model and its extended version captures endogeneity by employing an instrumental variable approach. The non-parametric model uses kernel regressions.

4.4.3 Theoretical issues in cost-minimization data envelopment analysis

Besstremyannaya and Simm (2019) propose a bootstrap methodology for estimating cost efficiency in data envelopment analysis. The paper considers the conventional concept of Färe et al. (1985) cost efficiency, for which the proposed algorithm re-samples "naive" input-oriented efficiency scores, uses them to rescale original inputs in order to bring them to the frontier, and then re-estimates cost efficiency scores for the rescaled inputs. Next, the paper examines Tone (2002) cost efficiency, where input prices vary across producers. Here Besstremyannaya and Simm (2019) show that the direct modification on bootstrap algorithms by Simar and Wilson (2007, 1998) is applicable.

The bootstrap methodology exploits these assumptions: 1) the sample are i.i.d. random variables with the continuous joint probability density function with support over production set; 2) monotonicity of technology, requirement of inputs for production ("no free lunch" condition), closedness and strict convexity of the production set, smoothness of the frontier; and 3) the probability of observing firms on the frontier approaches unity with an increase in sample. The analysis considers cases both with the absence and presence of environmental variables (i.e. input variables not directly controlled by firms). The results of simulations for a multi-input, multi-output Cobb–Douglas production function with correlated outputs, and correlated technical and cost efficiency, show consistency of the proposed algorithm (i.e. in terms of the coverage of the true confidence intervals of the estimate), even for small samples.

4.5 Policy evaluation

4.5.1 Average treatment effect

Besstremyannaya (2013) provides an example of regression analysis which evaluates the average treatment effect of the inpatient prospective payment system in Japan through difference-in-difference estimations. The analysis is applied to the values of cost efficiency or technical efficiency scores, computed according to parametric and non-parametric analysis. The treated group is acute-care Japanese local public hospitals which introduced the inpatient PPS in 2006. The control group are those acute-care

Japanese local public hospitals, which satisfy the criteria for participation in the PPS reform but remained under the traditional fee-for-service remuneration in 2006–2009. The paper follows the Dafny and Dranove (2006) approach to define the dependent variable as the difference between average values of each efficiency score in (three) pre-reform and (from one to three) post-reform years. The methodology goes in line with Bertrand et al. (2004) recommendation of collapsing the data into the pre-reform and post-reform periods to avoid inconsistency of the standard errors for the estimated coefficient for the reform effect.

Two approaches are used to estimate the average treatment effect: 1) mean unconditional comparison (where only the reform dummy enters the list of regressors), 2) mean conditional comparison (hospital variables are included in the list of regressors). Hospital heterogeneity with respect to the average treatment effect of the reform is interpreted as the significance of the estimated coefficients for hospital variables in the second approach. Although the values for these coefficients are not reported in the paper, the values of the average treatment effect differ under the first and second approach, which implies that (at least some of) hospital variables are significant.

4.5.2 Average treatment effect and conditional average treatment effect in the finite mixture models

Besstremyannaya (2015a) computed the treatment effect of the rise in the nominal coinsurance rate for heads of households in the non-national health insurance plans in 2003 in Japan. The study deals with consumers in each of the most probable latent class, according to the posterior analysis after the estimation of the loglinear and generalized finite mixture models of healthcare expenditure. The dependent variable is the difference in the fitted value of the healthcare expenditure in the pre-reform and post-reform years. In computation of the treatment effect in each latent class, the paper follows Hirano et al. (2000) methodology on measuring treatment effects for subpopulations. The treated group are those who experienced the rise in the nominal coinsurance rate and the control group are other respondents. The estimation of conditional average treatment effect uses income, age, education, self-reported health condition and the dummy for urban residence as variables for matching the treated and the controls in the regression analysis.[20]

The below approach is used for estimations. Firstly, the class-specific average treatment effect $\bar{\tau}_j$ is calculated as:

$$\bar{\tau}_j = \frac{1}{N}\sum_{i=1}^{N} E[d_{ij}(w_i = 1) - d_{ij}(w_i = 0)|\theta_j], \tag{20}$$

where d_{ij} is the difference in the fitted values of the dependent variable in the post- and pre-reform years for individual i in class j , w_i is the treatment indicator. It is estimated under assumption that all observations belong to the class j.

Secondly, Hirano et al. (2000) methodology is used to compute average treatment effect for each class $\bar{\tau}_j^{ATE}$:

$$\bar{\tau}_j^{ATE} = \frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{j=1}^{C} P(i \in j)\bar{\tau}_j, \tag{21}$$

[20] The dataset of the survey used for estimations includes only young and middle-aged adult respondents, 24–50 ears old, so the issues related to the special features of healthcare consumption by the elderly need not to be included in the analysis.

where $i = 1,...,N_j$ indicates individuals in class j, $P(i \in j)$ is posterior probability of class membership.

Linear estimator of the conditional average treatment effect is calculated as follows.

1. Obtain class-specific estimates for each j:

$$d_{ij} = \tau_j w_i + \kappa_{ij} h_i^{pre} + \psi_i, E\psi_i = 0, \tag{22}$$

where h^{pre} denotes the average values of covariates x (excluding price) in the pre-reform years, and the fitted value of τ_j give the linear estimate of the conditional average treatment effect in the class.

2. Weight the estimates by $P(i \in j)$ and average over subsamples.

To account for nonlinear effects, Besstremyannaya (2015a) also computes the conditional average treatment effect in matching and regression by averaging over sample and posterior distribution of covariates:

$$\overline{\tau(x)}_j^{CATE} = \frac{1}{N_j} \sum_{i=1}^{N_j} \sum_{j=1}^{C} P(i \in j) \left(E[d_{ij}(w_i = 1) - d_{ij}(w_i = 0) | h_i^{pre}, \theta_j] \right) \tag{23}$$

where θ_j is the vector of unknown parameters associated with class membership, see equation (1) in Section 3.1 of this summary.

The analysis touches upon the issues related to the non-overlap and unconfoundedness assumptions.

4.5.3 Treatment effect in dynamic panel data models

Dynamic panel data models are well-suited for evaluating heterogeneous effect of a binary (or a continuous) variable, which may be regarded as a policy indicator (or a measure of policy intensity). Two approaches are offered in this dissertation.

Besstremyannaya and Golovan (2022c) use the fact that the interaction term of the $y_{it-1} - x'_{it-1}\beta - \mu$ (and of corresponding terms for second order lag) and the reform is included in the list of variables in the right-hand of the equation (18). Specifically, the analysis of heterogeneity focuses on the values of the quality measures in the previous period(s) in the deciles of the conditional distribution of quality.

Besstremyannaya and Golovan (2022a) compute the long-term mean μ in equation (19) as a function of the lags of the dependent variable. Difference in $\mu(\alpha_t)$ and $\mu(\alpha_{t-1})$ enables to evaluate the impact of the reform cleansed of mean reversion, and the fact that μ depends on y_{t-1} allows to estimate groupwise effect of the reform. So the paper assesses how the quality incentive affects the aggregate quality measure at quintiles of the US Medicare hospitals.

5 Contribution

The novelty of the papers in this dissertation deal with the development of econometric models as well as with the application of the models to novel research tasks in order to reveal heterogeneity in economic choices by producers and consumers, to disentangle heterogeneous effects of exogenous shocks on firm costs, and to evaluate heterogeneous effects of policy reforms aimed at price and quality regulation.

The novel **theoretical results** of the main part of the dissertation are as follows.

1. The development of an approach for the bias-correction of the data envelopment analysis scores in the cost-minimization problems of Färe et al. (1985) and Tone (2002), see Besstremyannaya and Simm (2019).[21]
2. The discovery of the limited applicability of the Canay (2011) conditional quantile regression estimator with quantile-independent fixed effects for the analysis of short panels owing to the asymptotic bias (Besstremyannaya and Golovan, 2019, 2021).[22]
3. The development of an approach for accounting for the multivariate dependence of the policy variable in dynamic panel data models and disentangling two sources of intertemporal dependence: the policy effect and the impact of regression towards the mean (Besstremyannaya and Golovan, 2022a).

The novelty of the empirical papers in the main part of the dissertation consists in the novel application of theoretical methodology and in the econometric analysis of novel topics on agent heterogeneity and heterogeneous effect of reforms in various economics fields.

The novel use of theoretical methodology are

1. the study of the applicability and adaptation of generalized finite mixture models to analyze heterogeneity in consumer healthcare expenditure (Besstremyannaya, 2015a);
2. the modification of treatment effect estimators within finite mixture models to analyze the average treatment effect of a price reform on consumer healthcare expenditure (Besstremyannaya, 2015a);
3. the modification of dynamic panel data models to study the heterogeneous effect of price and incentive regulation in healthcare (Besstremyannaya and Golovan, 2022a,c);
4. the adaptation of a panel data stochastic frontier model with latent classes to study the cost efficiency of hospitals as multi-output producers of healthcare (Besstremyannaya, 2011);[23]
5. the adaptation of a panel data conditional quantile model to study the cost efficiency of banks as multi-output producers (Besstremyannaya, 2017a);[24]
6. the modification of a panel data conditional quantile model to study the production efficiency of hospitals as multi-output producers (Besstremyannaya and Golovan, 2022b);
7. the use of parametric and non-parametric models to evaluate the effect of hospital financing reform in difference-in-difference estimations, with an application to Japanese hospitals (Besstremyannaya, 2013).[25]

Novel topics in the analysis of heterogeneity are

1. the estimation of the heterogeneous association between R&D-to-sales ratio and the growth of Japanese innovative manufacturing firms (Besstremyannaya et al., 2022);
2. the estimation of the heterogeneous effects of the global financial crisis and the Great East Japan

[21] The working paper versions of the article and the code are cited by Hayashi (2017) in Tone's (2017) Handbook: *Advances in DEA Theory and Applications* and noted in review articles in the *Journal of Economic Surveys* (Daraio et al., 2019) and in the *Journal of Statistical Software* (Álvarez et al., 2020).

[22] The discussion of the Canay (2011) estimator, started in *Econometrics Journal* by Besstremyannaya and Golovan (2019), was followed by Chen and Huo (2021), who cite theoretical results and a meta-review of the literature in Besstremyannaya and Golovan (2019).

[23] Citations to the paper include the Greene's (2014) course "Stochastic Frontier Models and Efficiency Estimation" https://pages.stern.nyu.edu/~wgreene/FrontierModels.htm and Sickles et al. (2022) chapter on the applications of SFA to health economics https://economics.uq.edu.au/files/35634/WP052022.pdf

[24] References to the paper may be found in meta-reviews such as de Abreu et al. (2019) and Fukuyama et al. (2018).

[25] The paper is cited in meta-reviews, e.g. Emrouznejad and Yang (2018) and in numerous applications.

Earthquake of March 2011 on the cost efficiency of Japanese banks (Besstremyannaya, 2017a);

3. the estimation of the heterogeneous impact of the quality incentives in the US Medicare on the aggregate quality measure and its components with dynamic panel data models (Besstremyannaya and Golovan, 2022a,c);
4. the estimation of heterogeneity in the cost efficiency of Japanese hospitals and its interrelation with unobserved managerial practices (Besstremyannaya, 2011);
5. the estimation of the heterogeneous effect of a financial reform (prospective payment system) on technical and cost efficiency, and the average length of stay at Japanese hospitals (Besstremyannaya, 2013, 2016);
6. the estimation of the heterogeneous effect of the price reform (change in coinsurance rates) on healthcare expenditure through linear and generalized finite mixture models, with an application to Japanese consumer data (Besstremyannaya, 2015a);
7. the estimation of heterogeneity in income equity with respect to healthcare use and expenditure in Japan through latent class models: binary choice model, linear and generalized linear models (Besstremyannaya, 2017b);
8. the estimation of the heterogeneous effect of the type of health insurer (i.e. private or public) on the effectiveness/quality of regional healthcare systems in Russia (Besstremyannaya, 2015b).

6 Main findings

Heterogeneity in banking, with an application to Japan

1. There is technological heterogeneity in Japanese banking. According to the results of statistical tests, there is a more efficient path (low-cost quantiles) and a less efficient path (high-cost quantiles), so the effect of non-performing loans, non-traditional activities and bank profitability differs across high-cost and low-cost banks. Japanese banks demonstrate an inverse relationship between risk factors (e.g. the share of loan loss provisions in total loans), economies of scale and cost inefficiencies. Low-cost and high-cost banks show different associations between costs and risk-taking behavior (proxied by equity capital), the bank business model (proxied by an index of product diversity), and the regional macroeconomic environment. Business growth from economies of scale has a different association with credit risk (loan loss provisions or liquidity), profitability, and the business model (proxied by securities-to-loan ratio) at low-cost and high-cost banks (Besstremyannaya, 2017a).
2. There was a heterogeneous impact of the global financial crisis of 2007–2009 and of the Great East Japan Earthquake of March 2011 on costs, economies of scale and the cost inefficiency of Japanese banks. The effect of these two exogenous shocks and their time profiles differ across high-cost and low-cost banks. Such differences are argued to be inherent in the special features of bank profitability in Japan and the social role of banks (Besstremyannaya, 2017a).
3. There are differences (i.e. heterogeneity by Japanese bank charter) in the bias of the naive estimate of cost-efficiency according to cost-minimization DEA of Färe et al. (1985) or Tone (2002), see

(Besstremyannaya and Simm, 2019, 2015).[26]

Heterogeneity in the growth of innovative companies, with examples from Japan

1. There is a heterogeneous association between R&D-to-sales ratio and the growth of faster and slower growing Japanese innovative firms in each of the four manufacturing industries: chemicals and allied products; electronic and other electrical equipment; industrial and commercial machinery and computer equipment; and transportation equipment. There are statistical differences in the estimated coefficients for R&D intensity across low-, median- and high-growth firms within each industry (Besstremyannaya et al., 2022).
2. The association between R&D intensity and growth is strongest in two of Japan's four highly innovative industries: transportation equipment, and electronic and other electrical equipment. Moreover, the association between R&D-to-sales ratio and the growth of Japanese innovative manufacturing firms differs across pairs of industries. So strategies for firm growth in Japan require a degree of nuance. Specifically, R&D expenditure is vital for sustaining fast growth for firms in high-tech industries, but it may not be an engine of growth for slower-growing firms in less technology-intensive industries (Besstremyannaya et al., 2022).
3. Only in the group of high- and median-growth Japanese firms, do small firms grow faster than large firms. The effect of firm age on growth is negative only in the top quantiles and quantiles close to the median, while it is positive in the bottom quantiles. The stylized fact of the mean regression analysis, by which young firms grow faster than old ones, does not hold for slow-growing firms (Besstremyannaya et al., 2022).

Heterogeneity among producers and consumers of healthcare and heterogeneous effects of regulatory reforms, with applications to the US, Japan, and Russia

1. There is a direct association between prior quality (proxied by the aggregate quality measure) and the quality improvement owing to the incentive reform in US acute-care Medicare hospitals. The stylized fact in the prior literature, which states that a pay-for-performance incentive leads to greater improvements at hospitals with lower baseline quality needs to be reconsidered (Besstremyannaya and Golovan, 2022c).
2. There is a deterioration of the values for specific quality measures, which may be linked to the patient's benefit and hence to provider altruism at the highest-quality acute-care Medicare hospitals with respect to these measures (i.e. the communication of patients with medical personnel and the ability to receive help promptly). Other quality measures, less associated with patient benefits (e.g. of the clinical process of care) do not fall among the highest-quality hospitals. So there is heterogeneity of the effect of the incentive contract for the quality dimensions of altruistic providers (Besstremyannaya and Golovan, 2022c).
3. Japanese acute-care local public hospitals can be separated into two latent classes as regards their cost efficiency. The posterior probability of belonging to a more efficient class (in terms of lower

[26] "Heterogeneity depends on bank charters in the model with an intermediation approach: the distance from the 45 degree line is largest for national banks and long-term credit/trust banks. The bias and heterogeneity is larger in presence of the environmental variables." (Besstremyannaya and Simm, 2015, P.18)

costs) is associated with better values of three financial proxies for managerial practices: the ordinary balance ratio (the share of medical revenues in medical expenses), the share of transfers in medical revenues, and the share of labor costs in medical revenues (Besstremyannaya, 2011).

4. There is heterogeneity in productivity across Japanese acute-care local public hospitals with high and low output. There is a more efficient production path (high-output quantiles) and a less efficient production path (low-output quantiles), and there is a statistical difference in the values of input elasticities, input productivities, and the partial effects of hospital variables (i.e. hospital accreditation, the status of a designated hospital, and teaching activity) between high- and low-output quantiles. High-output hospitals show higher productivity by technicians, administrators and other staff, but lower productivity by physicians. High-output hospitals demonstrate better values for many indicators of managerial performance, which supports the idea that management and production are interrelated. The results point to an inexpedient mix of labor/capital and labor/medicines in all quantiles of hospital output, suggesting substantial opportunities for cost savings (Besstremyannaya and Golovan, 2022b).
5. There is a heterogeneous effect of the Japanese variant of the inpatient prospective payment system on the cost efficiency of acute-care hospitals, proxied by the average length of stay. Specifically, the length of stay goes up in the group of hospitals in the lowest percentiles of the pre-reform length of stay. There is heterogeneity in the effect on the quality of hospital care, proxied by the early readmission rate (Besstremyannaya, 2016). The effect of the inpatient prospective payment system on parametric and non-parametric efficiency scores of Japanese acute-care local public hospitals is limited and heterogeneous (Besstremyannaya, 2013). The findings point to inadequate incentives within the payment schedule (Besstremyannaya, 2016).
6. Japanese consumers (adults in Besstremyannaya (2017b) and young and middle-aged women in Besstremyannaya (2015a)) separate into latent classes with high and low healthcare expenditure and the posterior probability of class membership may be explained by health and lifestyle variables. The effect of price (the coefficient for the coinsurance rate) on healthcare expenditure by young and middle-aged women in Besstremyannaya (2015a) is negative and varies across classes. The effect is smaller (in absolute terms) among low users of healthcare, so the healthcare expenditure of these consumers is less price elastic. The values of each of the three estimators: average treatment effect, the effect in the linear estimations conditional on covariates and the conditional average treatment effect in matching and regression differ across the classes, which implies heterogeneity in the effect of the nominal coinsurance rate on healthcare expenditure (for high users and low users of healthcare). The values of the conditional average treatment effect estimator differ from the values of the average treatment effect estimator and the linear estimators, and this may be interpreted as the heterogeneity of the effect as regards consumer characteristics. The fact also highlights the importance of using a matched control group in the analysis.
7. The Japanese social insurance system is "pro-poor" as regards the use of outpatient or inpatient healthcare. The coefficients for the low income quintile (which approximates the poverty line in high-income countries under the OECD methodology) are significant in each of the latent classes. Regarding the income equity of consumer healthcare expenditure, the results reveal that the utilization of outpatient care is equitable in Japan with respect to disposable income. Concerning

outpatient or inpatient healthcare expenditure, Japanese adult consumers separate into three latent classes. Class membership is explained by using such proxies for lifestyle variables as indices of psychological distress and unhealthy habits (smoking and drinking) (Besstremyannaya, 2017b).

8. There is a positive and significant impact of private insurers on the quality of mandatory health insurance systems of Russian regions (Besstremyannaya, 2015b).

7 Theoretical and practical importance

7.1 Theoretical importance

The theoretical importance of the dissertation consists in the development of econometric theory and in the novel use of applied regression analysis. Firstly, Besstremyannaya and Simm (2019) pointed to the asymptotic bias of the naive scores in cost-minimization DEA and suggested a methodology for bias correction. Secondly, Besstremyannaya and Golovan (2019) discovered the non-applicability of the simple two-step estimator in the conditional quantile model with quantile-independent fixed effects by Canay (2011) for short panels and offered recommendations for practitioners. Besstremyannaya and Golovan (2021) further reviewed the modern techniques for estimating conditional quantile regressions in the case of short panels. Thirdly, an approach for estimating the policy effect of reform cleansed of mean reversion in dynamic panels is proposed in Besstremyannaya and Golovan (2022a).

The theoretical importance of the empirical analysis is the introduction of modern methods for studying heterogeneity to several economic fields: banking, the economics of innovation, health economics, and policy evaluation, and in the extensions of the dissertation, empirical finance, environmental economics, labor economics and municipal finance. The methods enable us to account for the observed and unobserved heterogeneity of producers and consumers and to disentangle the heterogeneous (and non-linear) effects of policy reforms.

The theoretical and empirical results of this dissertation are noted in handbooks and review articles in international journals.

7.2 Practical importance

The theoretical and empirical models developed in the dissertation were used in teaching the BA, MA and PhD students at the Department of Applied Economics and at the School of Finance of the Faculty of Economic Sciences of the HSE in 2011–2014 and in 2019–2023. This includes a PhD-level course “Modeling heterogeneity of economic agents”, MA-level courses/seminars on banking, behavioral finance, municipal finance, technological growth through innovation, applied health economics and a BA-level course “Economic growth and development”.

A large part of the papers of the dissertation was written within scientific projects of the HSE University International Laboratory for Macroeconomic Analysis in 2019–2023, and was financed the HSE University Basic Research Program. The results of other papers of the dissertation were employed in applied projects by NES/CEFIR in 2010–2019 (which include recommendations on Russian health reform for the Federal Mandatory Health Insurance Fund, a meta-review Besstremyannaya (2011) on calibrating the general equilibrium models for the Russian Ministry of Economy [27] and a textbook (Besstremyannaya, 2013, editor) on the budgetary process for the Russian Ministry of Finance),[28] by the Institute of Far Eastern Studies of the Russian Academy of Sciences in 2019, by the Central Economics and Mathematics Institute of the Russian Academy of Sciences in 2006–2019, and by the Center for Strategic Research in 2006–2007.

3 sole-authored and 1 co-authored policy briefs stemming from the research in the dissertation were published in 2013–2019 in The Forum for Research on Eastern Europe and Emerging Economies (FREE Network) series. They deal with comparative research on the urgent problems of Russian economy: innovation, municipal finance and health reforms.[29]

8 Approbation of the results of the research

The results of the theoretical and empirical parts of the dissertation were presented by the author at over 60 international conferences, workshops and invited lectures/seminars in the US, Europe and Japan. The conferences include the Annual Congresses by the European Economic Association and the European Meetings of the Econometric Society in 2013–2022; the Asian (2021, 2022), the Australasian (2013) and the North American Summer Meetings (2018, 2022) of the Econometric Society; the World Congress of the Econometric Society (2015); the Australasian Workshop on Econometrics and Health Economics (2010), the International Health Economics Association Congresses (2011, 2013), the Conferences of the American Society of Health Economists (2012, 2014); the HSE April conference (2021, 2022), Russian Economic Congress (2020).

A series of seminars by the author of the dissertation at the HSE Corporate Finance Center in 2019–2020 was devoted to the applicability of theoretical and empirical results on agent heterogeneity for teaching and research in corporate finance.

The robustness of the results of the research was investigated with the help of the following techniques: use of parametric and non-parametric models (Besstremyannaya, 2013, 2015b); employing alternate lists of inputs and outputs in productivity analysis, e.g. intermediation approach and asset approach in banking in Besstremyannaya and Simm (2019), different proxies for inpatient and outpatient activity of hospitals in Besstremyannaya (2013); comparison of various criteria in the analysis of the goodness-of-fit (Besstremyannaya, 2015a, 2017b); cross-validation and use of subsamples (Besstremyannaya, 2015a, 2016, Besstremyannaya and Golovan, 2022c); simulation analysis (Besstremyannaya and Simm, 2019, Besstremyannaya and Golovan, 2019, 2021); estimation of the treatment effects of the reform for varying number of post-reform years (Besstremyannaya, 2013, 2015a, 2016, Besstremyannaya and Golovan, 2022a); use of the results of extensive meta-reviews of the literature and methodology (Besstremyannaya, 2011, 2013, 2017b).

The international conferences, workshops and seminars in the past 5 years are listed below.

[27] http://old.economy.gov.ru/minec/activity/sections/NIR_NIOKR/doc20110406_02

[28] S.Strem, E.Leontyeva, edited by G.Besstremyannaya. Budgetary Process as an Instrument for Effective Governance. Ministry of Finance, Federal Treasury, International Bank for Reconstruction and Development. Moscow, 2013, ISBN 978-5-9710-0541-4

[29] https://freepolicybriefs.org/experts/galina-besstremyannaya/

1. The 2018 North American Summer Meeting of the Econometric Society (University of California, Davis).
2. The 2018 Annual Congress of the European Economic Association (University of Cologne).
3. The 2018 European Meeting of the Econometric Society (University of Cologne).
4. CINCH - Health Economics Research Center at University of Duisburg-Essen, Essen Health Economics Seminar, February, 2019.
5. The 12th Annual Conference on Innovation Economics (Northwestern University, 2019).
6. The 2019 Annual Congress of the European Economic Association (University of Manchester).
7. The 2019 European Meeting of the Econometric Society (University of Manchester).
8. The 2020 Annual Congress of the European Economic Association (virtual).
9. The Center for Econometrics and Business Analytics (CEBA) invited talk, September 2020 (virtual).
10. The 2020 Russian Economic Congress (virtual).
11. The 2021 Annual Congress of the European Economic Association (virtual).
12. The 2021 Asian Meeting of the Econometric Society (virtual).
13. The 2021 International Conference on Econometrics and Business Analytics (virtual).
14. The 2021 HSE April Conference, the 3rd International Workshop on Applied Econometrics (virtual).
15. The 2021 International Conference “Modern Econometric Tools and Applications” (virtual).
16. The 2021 Congress of the European Economic Association (virtual).
17. Waseda University, Waseda Institute of Political Economy Empirical Microeconomics Seminar, November 2021 (virtual).
18. The HSE University International Laboratory for Experimental and Behavioral Economics, online seminar, November 2021.
19. The HSE University 2022 April Conference, the 4th International Workshop on Applied Econometrics (virtual).
20. The 2022 North American Summer Meeting of the Econometric Society (University of Miami).
21. The 2022 Asian Meeting of the Econometric Society (China, virtual).
22. The 2022 Asian Meeting of the Econometric Society (Tokyo, virtual).
23. The 2022 Congress of the European Economic Association (Bocconi University).
24. The 2022 International Conference on Econometrics and Business Analytics (American University of Armenia).
25. The 2022 International Conference “Modern Econometric Tools and Applications” (the HSE University at Nizhny Novgorod).

9 List of author’s original articles

9.1 Main articles

Main results of the dissertation are published in the 14 physical articles (20 articles under the double weight of articles in the HSE A list or Web of Science/Scopus Q1 journal list). The articles constitute

19.5 physical author lists, of which Besstremyannaya's contribution is 16.5 author lists.

Besstremyannaya, G. (2011). Managerial performance and cost efficiency of Japanese local public hospitals: A latent class stochastic frontier model. *Health Economics*, 20(S1):19–34. (1 author list, the HSE A list, Scopus Q1, Web of Science Q1).

Besstremyannaya, G. (2013). The impact of Japanese hospital financing reform on hospital efficiency: A difference-in-difference approach. *The Japanese Economic Review*, 64(3):337–362. (1.5 author list, Scopus Q3, Web of Science Q4).

Besstremyannaya, G. (2015a). Heterogeneous effect of coinsurance rate on healthcare expenditure: Generalized finite mixtures and matching estimators. *Applied Economics*, 47(58):6331–6361. (1.75 author list, the HSE A list, Scopus Q2, Web of Science Q3).

Besstremyannaya, G. (2015b). Measuring the effect of health insurance companies on the quality of healthcare systems with kernel and parametric regressions (In Russian). *Applied Econometrics*, 38(2):3–20. (1 author list, Scopus Q4).

Besstremyannaya, G. (2016). Differential effects of declining rates in a per diem payment system. *Health Economics*, 25(12):1599–1618. (1 author list, the HSE A list, Scopus Q1, Web of Science Q1).

Besstremyannaya, G. (2017a). Heterogeneous effect of the global financial crisis and the Great East Japan Earthquake on costs of Japanese banks. *Journal of Empirical Finance*, 42:66–89. (1.5 author list, the HSE A list, Scopus Q1, Web of Science Q3).

Besstremyannaya, G. (2017b). Measuring income equity in the demand for healthcare with finite mixture models. *Applied Econometrics*, 46(2):5–29. (1 author list, Scopus Q4).

Besstremyannaya, G., Dasher, R., and Golovan, S. (2022). Quantifying heterogeneity in the relationship between R&D intensity and growth at innovative Japanese firms: A quantile regression approach. *Applied Econometrics*, 67:27–45. (1 author list, Besstremyannaya: 0.75 author lists, Scopus Q3).

Besstremyannaya, G. and Golovan, S. (2019). Reconsideration of a simple approach to quantile regression for panel data. *The Econometrics Journal*, 22(3):292–308. (2 author lists, Besstremyannaya: 1 author list, sections 1,4,5,S3, the HSE A list, Scopus Q1, Web of Science Q2).

Besstremyannaya, G. and Golovan, S. (2021). Measuring heterogeneity with fixed effect quantile regression: Long panels and short panels. *Applied Econometrics*, 64:70–82. (0.75 author lists, Besstremyannaya: 0.5 author lists, Scopus Q3).

Besstremyannaya, G. and Golovan, S. (2022a). Disentangling the impact of mean reversion in estimating policy response with dynamic panels. *Dependence Modeling*, 10(1):58–86. (2 author lists, Besstremyannaya: 1.5 author lists, Scopus Q3).

Besstremyannaya, G. and Golovan, S. (2022b). Measuring heterogeneity in hospital productivity: a quantile regression approach. *Journal of Productivity Analysis*, pages 1–29. available at https://link.springer.com/article/10.1007/s11123-022-00650-3, (2 author lists, Besstremyannaya: 1.75 author lists, the HSE A list, Scopus Q2, Web of Science Q2).

Besstremyannaya, G. and Golovan, S. (2022c). Provider altruism in incentives contracts: Medicare's quality race. *HSE Economic Journal*, 26:375–403. (2 author lists, Besstremyannaya: 1.75 author lists, Scopus Q3).

Besstremyannaya, G. and Simm, J. (2019). Estimation of cost efficiency in non-parametric frontier models. *St Petersburg University Journal of Economic Studies*, 35(1):3–25. (1 author list, Besstremyannaya: 0.5 author lists, Scopus, no quartile, the HSE D list).

9.2 Other articles

Besstremyannaya, G. (2006). Unified Social Tax reform and shadow sector in healthcare and education (In Russian). *Voprosy Ekonomiki*, (6):107–119. (1 author list, Scopus, no quartile).

Besstremyannaya, G. (2009a). Increased public financing and health care outcomes in Russia. *Transition Studies Review*, 16(3):723–734. (0.75 author list, Scopus Q3).

Besstremyannaya, G. (2009b). Micro data assessment of Russian drug benefit monetization. *Journal of Health Organization and Management*, 23(5):465–476. (0.75 author list, Scopus Q2).

Besstremyannaya, G. (2019a). Informal taxes for the provision of public goods in Russian regions (In Russian). *Voprosy Ekonomiki*, (1):124–134. (0.75 author list, Scopus Q2).

Besstremyannaya, G. (2019b). Strategies for growth through mergers and acquisitions: evidence from Russian companies (In Russian). *Financial Journal*, (4):50–59. (0.75 author list, the HSE D list).

9.3 REPEC working papers

Besstremyannaya, G. (2015). The adverse effects of incentives regulation in health care: a comparative analysis with the U.S. and Japanese hospital data. Working Papers w0218, New Economic School (NES). https://ideas.repec.org/p/abo/neswpt/w0218.html.

Besstremyannaya, G., Dasher, R., and Golovan, S. (2019a). Growth through acquisition of innovations. Working Papers w0247, New Economic School (NES). https://ideas.repec.org/p/abo/neswpt/w0247.html.

Besstremyannaya, G., Dasher, R., and Golovan, S. (2019b). Technological change, energy, environment and economic growth in Japan. Ruhr Economic Papers 797, RWI - Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen. https://ideas.repec.org/p/zbw/rwirep/797.html.

Besstremyannaya, G. and Golovan, S. (2019). Physician's altruism in incentive contracts: Medicare's quality race. CINCH Working Paper Series 1903, Universitaet Duisburg-Essen, Competent in Competition and Health. https://ideas.repec.org/p/duh/wpaper/1903.html.

Besstremyannaya, G. and Golovan, S. (2022). Instrumental Variable Quantile Regression For Clustered Data. HSE Working papers WP BRP 255/EC/2022, National Research University Higher School of Economics. https://ideas.repec.org/p/hig/wpaper/255-ec-2022.html.

Besstremyannaya, G. and Simm, J. (2015). Robust non-parametric estimation of cost efficiency with an application to banking industry. Working Papers w0217, New Economic School (NES). https://ideas.repec.org/p/abo/neswpt/w0217.html.

Besstremyannaya, G., Simm, J., and Golovan, S. (2017). Robust estimation of cost efficiency in non-parametric frontier models. Working Papers w0244, Center for Economic and Financial Research at New Economic School. https://ideas.repec.org/p/cfr/cefirw/w0244.html.

10 References

Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, 97(457):284–292.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A., Drukker, D., Herr, J. L., and Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *The Stata journal*, 4(3):290–311.

Abadie, A. and Imbens, G. (2002). Simple and bias-corrected matching estimators for average treatment effects. Technical Report NBER Technical Working Paper 283.

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1):1–11.

Abadie, A. and Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2):781–807.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Acemoglu, D., Akcigit, U., Hanley, D., and Kerr, W. (2016). Transition to clean technology. *Journal of Political Economy*, 124(1):52–104.

Acemoglu, D., Autor, D. H., and Lyle, D. (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of Political Economy*, 112(3):497–551.

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1):21–37.

Álvarez, I. C., Barbero, J., and Zofío, J. L. (2020). A data envelopment analysis toolbox for MATLAB. *Journal of Statistical Software*, 95:1–49.

Andrews, D. W. (1988). Chi-square diagnostic tests for econometric models: Introduction and applications. *Journal of Econometrics*, 37(1):135–156.

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824–827.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. *The Economic Journal*, 114(494):C52–C83.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Angrist, J. D. and Pischke, J.-S. (2015). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58:277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Athey, S., Imbens, G. W., and Wager, S. (2018). Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(4):597–623.

Bago d'Uva, T. (2005). Latent class models for use of primary care: evidence from a British panel. *Health Economics*, 14(9):873–892.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data, 3rd ed.* Springer.

Baltagi, B. H., Bratberg, E., and Holmås, T. H. (2005). A panel data study of physicians' labor supply: the case of Norway. *Health Economics*, 14(10):1035–1045.

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9):1078–1092.

Barnett, A. G., van der Pols, J. C., and Dobson, A. J. (2004). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1):215–220.

Barro, R. J. (2013). Inflation and economic growth. *Annals of Economics & Finance*, 14(1):85–109.

Bassett, G. W. and Koenker, R. (2017). A quantile regression memoir. In *Handbook of Quantile Regression*, pages 3–5. Chapman and Hall/CRC.

Battese, G. E. and Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, 3(1):153–169.

Battese, G. E. and Corra, G. S. (1977). Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics*, 21(3):169–179.

Bazzi, S. and Clemens, M. A. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal: Macroeconomics*, 5(2):152–86.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Biørn, E., Hagen, T. P., Iversen, T., and Magnussen, J. (2010). How different are hospitals' responses to a financial reform? The impact on efficiency of activity-based financing. *Health Care Management Science*, 13(1):1–16.

Bleaney, M., Gemmell, N., and Kneller, R. (2001). Testing the endogenous growth model: public expenditure, taxation, and growth over the long run. *Canadian Journal of Economics*, 34(1):36–57.

Bloom, N., Lemos, R., Sadun, R., and van Reenen, J. (2020). Healthy business? Managerial education and management in health care. *Review of Economics and Statistics*, 102(3):506–517.

Bloom, N., Sadun, R., and Van Reenen, J. (2016). Management as a technology? Technical report, National Bureau of Economic Research, WP 22327.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, 122(4):1351–1408.

Bloom, N. and van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1):203–24.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Blundell, R. and Stoker, T. M. (2007). Models of aggregate economic relationships that account for heterogeneity. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, pages 4609–4666. Elsevier.

Borovičková, K. and Shimer, R. (2017). High wage workers work for high wage firms. Technical Report 24074, National Bureau of Economic Research.

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260.

Browning, M. and Carro, J. (2007). Heterogeneity and microeconometrics modelling. *Econometric Society Monographs*, 43:47–74.

Browning, M., Hansen, L. P., and Heckman, J. J. (1999). Micro data and general equilibrium models. In Taylor, J. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1, chapter 8, pages 543–633. Elsevier.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data, 2nd ed.* Cambridge University pPess.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, 14(3):368–386.

Caprio, G. and Honokan, P. (2014). Banking crises. In *The Oxford Handbook of Banking*, pages 700–720. OUP Oxford.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1):1–25.

Charnes, A., Cooper, W. W., Huang, Z. M., and Sun, D. B. (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *Journal of Econometrics*, 46(1-2):73–91.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444.

Chen, L. and Huo, Y. (2021). A simple estimator for quantile panel data models using smoothed quantile regressions. *The Econometrics Journal*, 24(2):247–263.

Chernozhukov, V. and Hansen, C. (2004). Instrumental variable quantile regression. Working Paper, available at http://www.mit.edu/∼vchern.

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chernozhukov, V. and Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2):491–525.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1):379–398.

Chernozhukov, V., Hansen, C., and Wüthrich, K. (2017). Instrumental variable quantile regression. In Koenker, R., Chernozhukov, V., He, X., and Peng, L., editors, *Handbook of Quantile Regression*, pages 119–143. Chapman and Hall/CRC.

Clogg, C. C. (1981). Latent structure models of mobility. *American Journal of Sociology*, 86(4):836–868.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis*. Springer Science & Business Media.

Compiani, G. and Kitamura, Y. (2016). Using mixtures in econometric models: a brief review and some new results. *The Econometrics Journal*, 19(3):C95–C127.

Conan Doyle, A. (1890). *The Sign of Four*. London: Spencer Blackett.

Cunha, F., Heckman, J. J., and Navarro-Lozano, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *NBER Working Paper*, (w11024).

Dafny, L. and Dranove, D. (2006). Regulatory exploitation and the market for corporate controls. NBER Working Paper 12438, National Bureau of Economic Research Cambridge, Mass., USA.

Daraio, C., Kerstens, K. H., Nepomuceno, T. C. C., and Sickles, R. (2019). Productivity and efficiency analysis software: an exploratory bibliographical survey of the options. *Journal of Economic Surveys*, 33(1):85–100.

de Abreu, E. S., Kimura, H., and Sobreiro, V. A. (2019). What is going on with studies on banking efficiency? *Research in International Business and Finance*, 47:195–219.

Deb, P. and Holmes, A. M. (2000). Estimates of use and costs of behavioural health care: a comparison of standard and finite mixture models. *Health Economics*, 9(6):475–489.

Deb, P. and Trivedi, P. K. (1997). Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics*, 12(3):313–336.

Deb, P. and Trivedi, P. K. (2002). The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics*, 21(4):601–625.

Dhaene, G. and Jochmans, K. (2015). Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies*, 82(3):991–1030.

Dias, D. A. and Marques, C. R. (2010). Using mean reversion as a measure of persistence. *Economic Modelling*, 27(1):262–273.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association*, 78(383):605–610.

Duan, N., Manning, W. G., Morris, C. N., and Newhouse, J. P. (1983). A comparison of alternative models for the demand for medical care. *Journal of Business & Economic Statistics*, 1(2):115–126.

Emrouznejad, A. and Yang, G.-l. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-economic planning sciences*, 61:4–8.

Everitt, B. and Hand, D. (1981). *Mixtures of Discrete Distributions*. Springer.

Färe, R., Grosskopf, S., and Lovell, C. K. (1985). *The Measurement of Efficiency of Production*. Springer Science & Business Media.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3):253–281.

Fetter, R. and Freeman, J. (1986). Diagnosis Related Groups: product line management within hospitals. *Academy of Management Review*, 11(1):41–54.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Fried, H. O., Lovell, C. K., Schmidt, S. S., Schmidt, S. S., et al. (2008). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.

Friedman, M. (1992). Do old fallacies ever die? *Journal of Economic Literature*, 30(4):2129–2132.

Fukuyama, H., Färe, R., and Weber, W. L. (2018). Valuing and ranking Japanese Banks: Application to Japan Post Bank and Mizuho Bank. *Journal of the Operational Research Society*, 69(12):2021–2033.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15:246–263.

Galton, F. and Dickson, H. (1886). Family likeness in stature. *Proceedings of the Royal Society of London*, 40:42–73.

Galvao, A. F. and Kato, K. (2016). Smoothed quantile regression for panel data. *Journal of Econometrics*, 193(1):92–112.

Gao, A., Lin, Z., and Na, C. F. (2009). Housing market dynamics: Evidence of mean reversion and downward rigidity. *Journal of Housing Economics*, 18(3):256–266.

Gaud, P., Jani, E., Hoesli, M., and Bender, A. (2005). The capital structure of Swiss companies: an empirical analysis using dynamic panel data. *European Financial Management*, 11(1):51–69.

Gauss, C. (1823). *Theoria combinationis observationum erroribus minimis obnoxiae*. Gottingae: Apud Henricum Dieterich.

Geroski, P. A., Machin, S. J., and Walters, C. F. (1997). Corporate growth and profitability. *The Journal of Industrial Economics*, 45(2):171–189.

Goddard, J., Molyneux, P., and Wilson, J. O. (2004). The profitability of European banks: a cross-sectional and dynamic panel analysis. *The Manchester School*, 72(3):363–381.

González, V. M. and González, F. (2012). Firm size and capital structure: Evidence using dynamic panel data. *Applied Economics*, 44(36):4745–4754.

Greene, W. (2002). Alternative panel data estimators for stochastic frontier models. Technical report, Working paper, Stern School of Business, New York University (September 2002). available at https://pages.stern.nyu.edu/ wgreene/panelfrontiers.pdf.

Greene, W. (2003). *Econometric Analysis, 5th ed.* Pearson Education.

Greene, W. (2007). *Limdep 9.0. Econometric Modeling Guide*. Econometric Software, Inc., Plainview, NY.

Greene, W. (2012). *Econometric Analysis, 7th ed.* Pearson Education.

Griliches, Z. (1996). The discovery of the residual: A historical note. *Journal of Economic Literature*, 34(3):1324–1330.

Hagenaars, J. A. and McCutcheon, A. L. (2002). *Applied Latent Class Analysis*. Cambridge University Press.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1):83–116.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton.

Harding, M. and Lamarche, C. (2016). Penalized quantile regression with semiparametric correlated effects: An application with heterogeneous preferences. *Journal of Applied Econometrics*.

Härdle, W. and Linton, O. (1994). Applied nonparametric methods. In Engle, R. F. and McFadden, D., editors, *Handbook of Econometrics*, volume 4, pages 2295–2339. Elsevier.

Hayashi, F. (2000). *Econometrics*. Princeton University Press.

Hayashi, M. (2017). Using DEA to analyze the efficiency of welfare offices and influencing factors: The case of Japan's municipal public assistance programs. In Tone, K., editor, *Advances in DEA theory and applications: With extensions to forecasting models*, pages 300–314. John Wiley & Sons.

Hayfield, T. and Racine, J. (2011). Package 'np'. nonparametric kernel smoothing methods for mixed data types. version 0.40-4.

Hayfield, T. and Racine, J. S. (2008). Nonparametric econometrics: The np package. *Journal of Statistical Software*, 27:1–32.

Heathcote, J., Storesletten, K., and Violante, G. L. (2009). Quantitative macroeconomics with heterogeneous households. *NBER Working Paper*, (w14768).

Heckman, J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy*, 109(4):673–748.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320.

Heckman, J. and Vytlacil, E. (2007). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, pages 4779–4874. Elsevier.

Hirano, K., Imbens, G. W., and Ridder, G. (2000). Efficient estimation of average treatment effects using the estimated propensity score. NBER Technical Working Paper 0251.

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. *Health Economics*, 17(10):1107–1128.

Hsiao, C., Li, Q., and Racine, J. S. (2007). A consistent model specification test with mixed discrete and continuous data. *Journal of Econometrics*, 140(2):802–826.

Hughes, J. P. and Mester, L. J. (2013). Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation*, 22(4):559–585.

Imbens, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review*, 93(2):126–132.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics*, 86(1):4–29.

Jacobs, R. (2001). Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *Health Care Management Science*, 4(2):103–115.

Kato, K., Galvao Jr, A. F., and Montes-Rojas, G. V. (2012). Asymptotics for panel quantile regression models with individual effects. *Journal of Econometrics*, 170(1):76–91.

Kirman, A. (2006). Heterogeneity in economics. *Journal of Economic Interaction and Coordination*, 1(1):89–117.

Knapp, M., Gart, A., and Chaudhry, M. (2006). The impact of mean reversion of bank profitability on post-merger performance in the banking industry. *Journal of Banking & Finance*, 30(12):3503–3517.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89.

Koenker, R. (2005). *Quantile Regression*. Cambrdige University Press.

Koenker, R. (2017). Quantile regression: 40 years on. *Annual Review of Economics*, 9:155–176.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, pages 33–50.

Koenker, R. and Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448):1296–1310.

Kooreman, P. (1994). Data envelopment analysis and parametric frontier estimation: complementary tools. *Journal of Health Economics*, 13(3):345–346.

Kyereboah-Coleman, A. (2008). Corporate governance and firm performance in Africa: A dynamic panel data analysis. *Studies in Economics and Econometrics*, 32(2):1–24.

Laeven, L., Levine, R., and Michalopoulos, S. (2015). Financial innovation and endogenous growth. *Journal of Financial Intermediation*, 24(1):1–24.

Laffont, J.-J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*. MIT press.

Li, H., Lindsay, B. G., and Waterman, R. P. (2003). Efficiency of projected score methods in rectangular array asymptotics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):191–208.

Li, Q. and Racine, J. (2003). Nonparametric estimation of distributions with categorical and continuous data. *Journal of Multivariate Analysis*, 86(2):266–292.

Li, Q. and Racine, J. S. (2008). Nonparametric estimation of conditional cdf and quantile functions with mixed categorical and continuous data. *Journal of Business & Economic Statistics*, 26(4):423–434.

Machado, J. A. F. and Santos Silva, J. M. C. (2019). Quantiles via moments. *Journal of Econometrics*, 213(1):145–173.

Machin, S. and Van Reenen, J. (1993). Profit margins and the business cycle: evidence from UK manufacturing firms. *The Journal of Industrial Economics*, pages 29–50.

Mark, B. A., Harless, D. W., McCue, M., and Xu, Y. (2004). A longitudinal examination of hospital registered nurse staffing and quality of care. *Health Services Research*, 39(2):279–300.

McCullagh, P. and Nelder, J. A. (2019). *Generalized Linear Models*. Routledge.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. John Wiley and Sons.

Meeusen, W. and van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, pages 435–444.

Michaud, P.-C. and van Soest, A. (2008). Health and wealth of elderly couples: Causality tests using dynamic panel data models. *Journal of Health Economics*, 27(5):1312–1325.

Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4):111–132.

Nakayama, N. (2003). A comparison of parametric and non-parametric distance functions: A case study of Japanese public hospitals (in Japanese). *Journal of Healthcare and Society*, 13(1):83–95.

Oliveira, B. and Fortunato, A. (2006). Firm growth and liquidity constraints: A dynamic analysis. *Small Business Economics*, 27(2):139–156.

Parente, P. M. and Santos Silva, J. M. (2016). Quantile regression with clustered data. *Journal of Econometric Methods*, 5(1):1.

Parmeter, C. F. and Racine, J. S. (2019). Nonparametric estimation and inference for panel data models. *Panel Data Econometrics*, pages 97–129.

Pérez-Calero, L., del Mar Villegas, M., and Barroso, C. (2016). A framework for board capital. *Corporate Governance*.

Phillips, R. F. (2003). Estimation of a stratified error-components model. *International Economic Review*, 44(2):501–521.

Quetelet, M. (1842). *A Tretise on Man and the Development of His Faculties*. Edinburgh: William and Robert Chambers.

Racine, J. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, 119(1):99–130.

Racine, J. S. (2019). *An Introduction to the Advanced Theory of Nonparametric Econometrics: a Replicable Approach using R*. Cambridge University Press.

Ray, S. C. (2004). *Data envelopment analysis: theory and techniques for economics and operations research*. Cambridge University Press.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics*, 71(1):135–158.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Santos, G. F. (2013). Fuel demand in Brazil in a dynamic panel data approach. *Energy Economics*, 36:229–240.

Schennach, S. M. (2020). Mismeasured and unobserved variables. In Durlauf, S. N., Hansen, L. P., Heckman, J. J., and Matzkin, R. L., editors, *Handbook of Econometrics*, volume 7, pages 487–565. Elsevier.

Seiford, L. M. (1996). Data envelopment analysis: the evolution of the state of the art (1978–1995). *Journal of Productivity Analysis*, 7(2):99–137.

Seiford, L. M. and Thrall, R. M. (1990). Recent developments in DEA: the mathematical programming approach to frontier analysis. *Journal of Econometrics*, 46(1-2):7–38.

Sen, A. (2004). Economic methodology: heterogeneity and relevance. *Social Research: An International Quarterly*, 71(3):583–614.

Shleifer, A. (1985). A theory of yardstick competition. *RAND Journal of Economics*, 16:319–327.

Sickles, R. C., Wang, Z., and Zelenyuk, V. (2022). Stochastic Frontier Analysis for Healthcare, with Illustrations in R. Technical Report WP05/2022,ISSN No.1932-4398, University of Queensland, Centre for Efficiency and Productivity Analysis.

Silva, J. M. S. and Windmeijer, F. (2001). Two-part multiple spell models for health care demand. *Journal of Econometrics*, 104(1):67–89.

Simar, L. and Wilson, P. (2008). Statistical inference in nonparametric frontier models: recent developments and perspectives. In O.Fried, H., Lovell, C., and Schmidt, S., editors, *The Measurement of Productive Efficiency and Productivity Growth*, pages 421–521. Oxford University Press.

Simar, L. and Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1):49–61.

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2):76–90.

Stigler, S. M. (1997). Regression towards the mean, historically considered. *Statistical Methods in Medical Research*, 6(2):103–114.

Street, A., O'Reilly, J., Ward, P., and Mason, A. (2011). DRG-based hospital payment and efficiency: Theory, evidence, and challenges. In Busse, R., Geissler, A., Quentin, W., and Wiley, M., editors, *Diagnosis related groups in Europe. Moving towards transparency, efficiency, and quality in hospitals*, pages 93–114. Open University Press.

Thanassoulis, E., Portela, M., and Despic, O. (2008). Banking. In O.Fried, H., Lovell, C., and Schmidt, S., editors, *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.

Tone, K. (2002). A strange case of the cost and allocative efficiencies in DEA. *Journal of the Operational Research Society*, 53(11):1225–1231.

Tone, K. (2017). *Advances in DEA theory and applications: With extensions to forecasting models*. John Wiley & Sons.

Tosi, H. L. and Gomez-Mejia, L. R. (1989). The decoupling of CEO pay and performance: An agency theory perspective. *Administrative Science Quarterly*, pages 169–189.

Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics*, 17(2):127–147.

van Reenen, J. (2011). Wage inequality, technology and trade: 21st century evidence. *Labour Economics*, 18(6):730–741.

Van Ryzin, J. and Wang, H. (1978). The methods for smooth estimation of discrete distributions. *Communications in Statistics-Theory and Methods*, 7(3):211–228.

Verbeek, M. (2004). *A Guide to Modern Econometrics, 2nd ed.* John Wiley & Sons.

Wang, M.-C. and Van Ryzin, J. (1981). A class of smooth estimators for discrete distributions. *Biometrika*, 68(1):301–309.

Wedel, M. and DeSarbo, W. (2002). Mixture regression models. In Hagenaars, J. A. and McCutcheon, A. L., editors, *Applied Latent Class Analysis*, pages 366–370. Cambridge University Press.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126:25–51.

Wooldridge, J. (2007). Quantile methods. NBER Summer Institute, Lecture Notes 14, available at http://www.nber.org/WNE/lect_14_quantile.pdf.

Wooldridge, J. (2011). Thoughts on heterogeneity in econometric models. Presidential address, Midwest Economics Association. March 19, 2011.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data, 2nd ed.* MIT press.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach, 5th ed.* South Western CENGAGE Learning.

Wu, J.-L. (2000). Mean reversion of the current account: evidence from the panel data unit-root test. *Economics Letters*, 66(2):215–222.

Zhao, M., Bazzoli, G. J., Clement, J. P., Lindrooth, R. C., Nolin, J. M., and Chukmaitov, A. S. (2008). Hospital staffing decisions: Does financial performance matter? *Inquiry*, 45(3):293–307.

11 Supplement: Extensions of the research

This section presents four types of extensions of the research which were published in additional articles to this dissertation and in REPEC working papers.

The first group are macroeconomic papers. They follow the modern approach which is denoted in *Handbook of Macroeconomics* as the need of macroeconomic theory to account for the fact that "the empirical microeconomics literature on consumption, saving and labor supply reveals quantitatively important heterogeneity in agent preferences, constraints, in dimensions of labor supply and skill, and in human capital accumulation processes" (Browning et al., 1999, P.5).

The second group of extensions concerns the development of econometric theory, and an illustration of its application to empirical labor economics. The third group deals with the use of early methods of accounting for heterogeneity (e.g. through a fixed effect model as in Besstremyannaya (2009a)) or by employing statistical methods of measuring the heterogeneous effect of a reform under the unavailability of the data for regression analysis.

The fourth group of extensions provides evidence on agent heterogeneity through correlation analysis and descriptive analysis, often applied to aggregated forms of secondary data. The descriptive analysis may be silent about the standard errors and the statistical significance of the estimates but nonetheless it may be helpful as the first step in discovering the "anecdotal evidence" on agent heterogeneity.

11.1 Macroeconomic theory and applications

Besstremyannaya (2006) and the working papers Besstremyannaya et al. (2019a) and Besstremyannaya et al. (2019b) use microeconomic evidence on the heterogeneity of economic agents in developing general equilibrium models. Specifically, Besstremyannaya (2006) considers a macro economic model of the Russian economy with official and unofficial sectors, and assumes that workers can move between industries and between official and unofficial sector. Besstremyannaya et al. (2019a) develop a general equilibrium model with several strategies of firm growth: various types of innovation and technological merger. Besstremyannaya et al. (2019b) employ the Acemoglu et al. (2016) macro economic model with several technologies for firm production (i.e. more or less environmentally friendly technologies) and calibrate it with data for the Japanese economy in order to forecast economic growth and consumer welfare in the medium- and long-run.

11.2 Econometric theory with an application to empirical labor economics

Besstremyannaya and Golovan (2022) extend the model in Chernozhukov and Hansen (2006) and enable inference for conditional quantile regressions with endogenous covariates and clustered data. The paper proves that the Chernozhukov and Hansen (2006) instrumental variable quantile regression estimator is consistent where there is correlation of errors within clusters, and derive the asymptotic distribution for the estimator. As regards inference based on the instrumental variable quantile regression process as a whole, the paper extends the methodology of Chernozhukov and Hansen (2006) which uses a bootstrap to compute critical values of the test statistics. Besstremyannaya and Golovan (2022) propose resampling by clusters and prove that it offers an approach to this computation and, hence, to the implementation of asymptotic tests. The theoretical results concerning the asymptotic properties of the instrumental variable quantile regression estimator for clustered data are supported by the simulation analysis.

As an empirical analysis extending the results of the dissertation on the applicability of a conditional quantile regression for measuring heterogeneity, Besstremyannaya and Golovan (2022) further use a conditional quantile regression under endogeneity to estimate earning equations of US men and women where the female labor supply is endogenous and subject to the shock of World War II. It is plausible to assume that the effect of labor supply differs across high-wage and low-wage workers. Specifically, a shortage of highly-skilled labor causes an increase of wages in that segment, which do not decline despite a subsequent increase of the supply of highly-skilled labor.[30] The paper estimates the quantile regression analogs of the Acemoglu et al. (2004) two-stage least squares models for wage equations for men and women, and data are clustered at the state-year level. The results demonstrate that with an increase in quantile index, the coefficient for female labor supply becomes smaller in absolute terms. So the effect is weaker for higher-wage workers. The effect is statistically insignificant at high and median values of τ. The failure to incorporate the clustered structure of the data leads to erroneous conclusions about the effect of female labor supply on the earnings of high-wage male and female workers.

11.3 Basic econometric and statistical methods in empirical corporate finance, municipal finance and health economics

Besstremyannaya et al. (2019a) use panel data OLS regression with firm-level and industry-level fixed effects to estimate the impact of innovation and technological acquisitions on growth of Japanese large and medium-sized firms in 1999–2013 (Table 1, PP.27–28).

Besstremyannaya (2019a) offers a descriptive illustration of computing the heterogeneous effect of a reform in municipal finance. Specifically, the average effect of the potential policy change (and the standard deviation of the effect) is computed to mimic the difference-in-difference estimations. The analysis focuses on the share of voluntary contributions by citizens in non-tax revenues of municipal (and hence consolidated regional) budgets in Russia. Regions may co-finance such contributions, and a rise in the coefficient for the regional co-financing of the voluntary contributions is regarded as an example of a policy reform. The paper uses an example of such reform in the Perm region in 2014, and only one region is regarded as a treatment group[31] to assess its average effect on groups of Russian regions. The paper considers that there are two control groups of Russian regions: one group has negligible values for the share, and it is assumed that the model with the decreasing elasticity of the growth of this share is applicable. The other group has higher values for this share and a model with the constant elasticity of growth is employed. The paper assesses the average effect of the reform separately in each of the two groups of regions.

Besstremyannaya (2009a) uses a panel data fixed effect model to estimate the impact of the rise in public financing on the effectiveness of the regional healthcare systems in Russia. It employs the first differences in infant mortality and maternal mortality as dependent variables, and these outcomes proxy the effectiveness of healthcare provision. Besstremyannaya (2015b) uses an instrumental variable

[30] See evidence for France, Germany, Austria, the US and the UK in respectively, Abowd et al. (1999), Andrews et al. (2012), Borovičková and Shimer (2017), van Reenen (2011).

[31] Hence, regression analysis becomes infeasible.

approach to account for endogeneity in the choice of the mandatory health insurance system by Russian regions: with or without private health insurers. The dependent variable in the second-stage equation is one of the aforementioned indicators of the effectiveness of regional healthcare systems.

11.4 Correlation analysis and descriptive analysis in energy economics, empirical corporate finance and health economics

Besstremyannaya et al. (2019b) focus on over 500 Japanese energy companies in 1989–2013 and examines heterogeneity in their productivity, interpreted as productivity differences of clean and dirty technologies. Following the Acemoglu et al. (2016) approach, the paper uses the patent data for each energy product[32] in Japan and computes technology gap as the difference between the number of patents for clean and dirty technologies (normalized by the total number of patents per corresponding product). The resulting distribution of technology gap in Japan is given on Figure 2 of the paper, while section 5.4. describes the methodology and contrasts the results for the US and Japan.

Besstremyannaya (2019b) complements the theoretical model of Besstremyannaya et al. (2019a) by focusing on country-level micro evidence on strategies for company growth. The analysis uses secondary and aggregated data on Russian manufacturing and service companies. The results hint at the interrelation between competition in the industry and company decisions about innovation and technological M&As in Russia.

Besstremyannaya (2009b) employs correlation analysis to study regional variation of the early effect of drug benefit monetization in Russia.

[32] Defined according to the main the US SIC 3-digit code of the patenting firm.