Heterogeneity in Economics and Aggregation

Inaugural–Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts– und Gesellschaftswissenschaften
durch die
Rechts– und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von
Michal Paluch
aus Opole, Polen

Bonn 2008

Dekan: Prof. Dr. Erik Theissen
Erstreferent: Prof. Dr. Werner Hildenbrand
Zweitreferent: Prof. Dr. Monika Merz

Tag der mündlichen Prüfung: 4. Juli 2008

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

to my parents

Acknowledgements

A number of people and institutions were integral to the completion of this dissertation. Foremost, I am indebted to Professor Werner Hildenbrand for arousing my interests in aggregation theory and his excellent and firm guidance throughout the last four years. I also owe special thanks to Professor Alois Kneip for his inexhaustible support and very valuable discussions and comments on empirical problems related to this thesis. I really enjoyed working with them on a joint paper and benefited extensively from their supervision. It was they who showed me how to approach an economic problem in a scientific way.

I would also like to express my gratitude to Professors Monika Merz and Zeno Enders for being members of the doctoral committee.

I am very grateful for financial support of Bonn Graduate School of Economics and its director Professor Urs Schweizer. Without this help my dream of doctoral studies would have never come true. It also allowed me to visit international workshops and conferences, which made our doctoral program even more appealing. Thanks go also to the European Commission for the Marie Curie Scholarship which financed my stay at the Universitat Pompeu Fabra in Barcelona. During this stay I received very intensive guidance from Professor Joan Esteban, whose support is gratefully appreciated.

Discussions and stimulating comments of many researchers and friends greatly improved this dissertation. I would like to sincerely thank the members of the Forschungsgruppe Hildenbrand, especially to Jürgen Arns, Kaushik Bhattacharya, Manisha Chakrabarty, and Anke Schmalenbach. Many thanks also go to Davide Fiaschi, Mario Lavezzi, and Grazia Pittau for their fruitful remarks and discussions. I owe great thanks to Steffen Altmann, Oliver Fries, Jürgen Gaul, Sotiris Georganas, Benjamin Huck, Simon Junker, Lars Koch, Marc Schiffbauer, David Schröder, Matthias Wibral, and Michael Zähringer not only for their support, kindness, and brainpower, but mostly for making Bonn, as rainy as it can be, the sunniest place on Earth.

I would also like to express my sincere gratitude to Jurek and Ulla Sawczuk for their unfailing help and encouragement during my studies. I am very thankful to Marta, because her humor, understanding, emotional support, and unknown volumes of inspiration has kept me happy over the years. Finally, I would like to thank to my parents, my brother, and my sister for their endless love and great moral support in everything I do. My gratitude to them is beyond any words.

Contents

List of Tables

List of Figures

Introduction

Heterogeneity is all around us. People differ substantially with respect to their tastes, expectations, and available resources, and firms are heterogeneous in technologies and the quantity or nature of input factors. Since macroeconomics deals, amongst others, with relationships between aggregates over heterogeneous populations of economic agents, sound macroeconomic models should not neglect nor implausibly restrict the variety of agents. In models without behavioral heterogeneity there is no room for distributional considerations, trade, asymmetric information or coordination of individuals. Finally, the presence of heterogeneity is crucial to guarantee the uniqueness and stability of equilibrium.

Yet, many macroeconomic models treat heterogeneity in a very simplistic and often trivial way. They are based on the notion of the representative agent (RA) who solves an explicitly stated optimization problem and whose choices coincide with the aggregate choices of the heterogeneous individuals or firms. The traditional aggregate consumption function and aggregate production function are best examples of this modeling approach. However, the logical consistency of these models, that is, the compatibility with microeconomic theories of behavior, requires very restrictive and implausible assumptions on the heterogeneity of the population *and* on the model of individual behavior. The straightforward conclusion is that traditional aggregate consumption and production functions, in general, do not exist. This negative theoretical result is far from being new, as its earliest version goes back to Antonelli (1886).

Surprisingly, despite these theoretical shortcomings, RA models are still common practice in the literature. It is surprising, since there exists a well developed statistical aggregation theory. This theory can be used to build aggregationally consistent macroeconomic models, which are flexible enough to include very general types of heterogeneity and analyze its effect on macroeconomic outcomes. In this context, Hildenbrand and Kneip (1999, 2005) proposed an approach, which concentrates directly on modeling economic aggregates in terms of entire distributions of individual variables. Its main advantage is that the parametric specification of the individual model of behavior and the pattern of heterogeneity in the population is not needed. The cornerstone of their model are distributional assumptions of the structural stability type, which are supported by empirical studies. Chapters 1 and 2 of this thesis are closely related to their work. In particular, in Chapter 1 a growth model for an economy consisting of firms which are heterogeneous in technologies and input demands is proposed. Aggregation over these firms is carried out according to the distributional approach established in Hildenbrand and Kneip (2005). It is shown that the growth rate in this economy depends not only on changes in the aggregate level of capital and labor, but also on changes in the allocation of these inputs across firms. As the latter effects are neglected in conventional growth models, they are misleadingly captured by the residual TFP measure. In contrast, one is able to quantify the influence of these components in our study. An empirical analysis, which is based on structural estimation from firm-level data, reveals that changes in allocation of capital and labor have pronounced effects on GDP-growth for most European countries. Further, cross-country differences in the distributional effects are taken into account to improve conventional growth accounting exercises. In particular, it is found that they explain a sizable part of growth differences across countries.

Chapter 2 analyzes the property of structural stability of the joint distribution of households' income and wealth, which is required to build an aggregatively consistent model for aggregate consumption of a heterogeneous population. Empirical analysis based on the U.K. Family Resources Survey data from the period 1996-2001 examines whether the sequence of these distribution is structurally stable in the sense related to Malinvaud (1993). Hence, the main objective of this chapter is to look for the local time-invariance of a distribution derived by applying simple transformations like scaling or standardizing to the original distribution. This analysis makes use of kernel density estimation to identify the changes in shapes of the aforementioned distributions and to perform a nonparametric density time-invariance test as proposed by Li (1996). The main result is that accounting only for the changes in the vector of means of the original distribution is not sufficient to obtain the desired local time-invariance. In fact, this can be achieved by accounting for changes in the vector of means *and* dispersion parameters of the original distribution.

Finally, Chapter 3 deals with different concepts of income elasticities of demand for a heterogeneous population and the relationship between individual and aggregate elasticities is analyzed. In particular, it allows to compare the income elasticity of demand of the representative agent with the distribution of individual income elasticities. It is shown that, in general, the aggregate elasticity is not equal to the mean of individual elasticities. The difference depends on the heterogeneity of the population and is quantified by a covariance term. Sign and magnitude of this term are determined by an empirical analysis based on the U.K. Family Expenditure Survey. It is shown that the relevant quantities can be identified from cross-section data and, without imposing restrictive structural assumptions, can be estimated by nonparametric techniques. It turns out that the aggregate elasticity significantly overestimates the mean of individual elasticities for many commodity groups.

Chapters 1 to 3 represent self-contained units.

Chapter 1

Distributional Effects of Capital and Labor on Economic Growth*

1.1 Introduction

In the following we propose a growth model for an economy consisting of firms which are heterogeneous in technologies and input demands. We show that the growth rate in this economy depends not only on changes in the aggregate level of capital and labor, but also on changes in the allocation of these inputs across firms. As the latter effects are neglected in conventional growth models, they are misleadingly captured by the residual measure, referred to as total factor productivity (TFP). In contrast, we are able to quantify the influence of these components by structural estimation from firm-level data. Further, we take cross-country differences in the distributional effects into account to improve conventional growth accounting exercises.

Why do some countries grow and others stagnate?[1] This question initiated the growth accounting literature, which assigns cross-country differences in growth or income to differences in physical and human capital as well as the unobservable efficiency with which input factors are combined. The consensus view in this literature is

*This chapter is based on Paluch and Schiffbauer (2007).

[1]The Science magazine considers this question as one of the 125 "most compelling puzzles and questions facing scientists today" (Science, 2005).

that only approximately one third of the cross-country growth or income differences is explained by differences in input factors. The remaining two thirds are left unexplained and attributed to differences in the unobservable efficiency which is referred to as total factor productivity (TFP).[2] In this context, Abramovitz (1956) refers to TFP as the *measure of our ignorance*.

The fact that TFP is unobservable and at the same time explains the major part of cross-country differences triggered tremendous efforts to identify its determinants in recent years.[3] However, we show in this chapter that the above mentioned growth accounting results have to be revised if one consistently aggregates over heterogeneous firms. In order to illustrate the relevance of aggregation for growth models we briefly discuss fundamental results of the aggregation literature.

The pillar of every macroeconomic growth model is an aggregate production function F, which relates aggregate capital $\bar{K}$ and labor $\bar{L}$ to aggregate output $\bar{Y}$, i.e., $\bar{Y} = F(\bar{K}, \bar{L})$. However, although there exists a well developed microeconomic theory of production for a single firm, there is no corresponding theoretical foundation for the entire economy. In fact, the aggregate production function suffers from two types of aggregation problems. The first, often referred to as the "measurement problem," involves the aggregation of different types of capital, labor, and output within a firm into one capital and labor input and one output. The second is concerned with aggregation of heterogeneous technologies and input demands across firms into their aggregate counterpart. These problems have been dealt with extensively in the aggregation literature. Early works by Nataf (1948), Gorman (1953), and a series of papers by Franklin Fisher (collected in Fisher, 1993)[4] have shown that in the absence of perfect competition and perfect factor mobility the aggregate production function F cannot be linked to microeconomic production functions unless all firms operate according to identical and constant returns to scale technologies.

[2] See, for example, Caselli (2005), Hall and Jones (1999) or Jorgensen (2005).

[3] This issue is best summarized by the title of a recent paper by Prescott (1998) "Needed: A Theory of Total Factor Productivity."

[4] For a comprehensive survey on aggregation of production functions, see Felipe and Fisher (2003).

A frequent short-cut that circumvents the problem of aggregation over heterogeneous technologies is the assumption that the production function of an entire economy complies with the one of a single *representative firm.* Although the above theoretical results show that this link is only possible under very restrictive assumption, it is often applied in theoretical and empirical analysis due to its simplicity. However, from a practical point of view, growth models that ignore consistent aggregation over heterogeneous firms will suffer from serious drawbacks:[5] they neglect growth effects of (i) changes in the allocation of inputs[6] and (ii) changes in the pattern of economic interactions between firms. Yet, it is reasonable to expect that these factors affect growth substantially, since they represent changes in growth due to changes in the market structure. For example, differences in the degree of competition in different industries as well as different incentives to innovate for small, medium, and large firms are found to affect technological change (see, e.g., Aghion and Griffith, 2005). Where are these effects in the growth literature? As they are not assigned to the levels of aggregate capital or labor, they are assigned to the unobserved efficiency. Therefore, they are misleadingly captured by the residual TFP measure.

In order to assess the impact of changes in the allocation of capital and labor on growth, we apply the aggregation procedure established by Hildenbrand and Kneip (2005). Our main result is that the growth rate of aggregate output depends on changes in the levels of aggregate capital and labor as well as changes in the distribution of capital and labor in the economy. We quantify the growth effect of each component by means of structural estimation based on firm-level data. These effects

---

[5]Hopenhayn (1992) initiated a literature on the effect of firm heterogeneity on industry dynamics. His approach was extended, e.g., by Melitz (2003) to analyze the impact of trade liberalization on the aggregate productivity of an economy. In these models firms are heterogeneous in productivity which is included in a way such that the impact of the productivity distribution on aggregate demand for inputs is fully determined by the average productivity. Consequently, under this parsimonious aggregation rule, aggregate output depends on average productivity and average input demands but not on the allocation of inputs across firms. That is, once the average productivity level is determined the model yields identical aggregate outcomes as a model based on a representative firm.

[6]Empirical studies document that these changes are substantial in developed and developing countries. For example, Roberts and Tybout (1997) quantify the rate of labor reallocation among manufacturing firms between 25 and 30 percent.

are estimated separately for each of 20 European countries. Our main findings are that distributional effects are significant in all countries. Further, they are as large as the corresponding level effects in most countries. Finally, we exploit the information on the different distributional changes across countries to conduct a growth accounting exercise. More precisely, we assess the explanatory power of the distributional changes with respect to cross-country growth differences. It turns out that these effects explain additionally up to 17%. Accordingly, an aggregation approach that consistently accounts for firm heterogeneity can help explain the growth path of a single country as well as cross-country growth differences. Hence, the role of capital and labor in explaining the growth path of a single country or growth differences across countries is understated if these aggregation issues are not taken into account.

In the next section, we present our growth model for an economy consisting of heterogeneous firms. In Section 3, we describe the data, the empirical strategy, and discuss our results. Section 4 presents the growth accounting exercise, whereas the final section concludes.

1.2 The Model

Assume that in period t each firm j from a heterogeneous population of firms J_t produces according to a firm-specific production function $f_t^j(\cdot)$ defined by

$$Y_t^j = f_t^j(K_t^j, L_t^j),$$

where Y_t^j denotes the output level, K_t^j the capital stock and L_t^j the labor demand.[7] Further, we assume that the heterogeneity in production functions f_t^j, i.e., in technologies and input demands, can be parametrized by a vector of parameters V_t^j. In general, V_t^j is unobservable. Then one can write

$$Y_t^j = f(K_t^j, L_t^j, V_t^j). \tag{1.1}$$

[7]The model can be extended to the case of multiple capital and labor inputs.

Hence, technological changes over time translate into changes in the distribution of V_t^j across J_t. The function f can therefore, without loss of generality, be regarded as time-invariant and equal for all firms. In the simplest scenario, f could be a Cobb-Douglas production function with $V_t^j = (V_{1,t}^j, V_{2,t}^j)$ such that $Y_t^j = V_{1,t}^j \cdot {K_t^j}^{V_{2,t}^j} \cdot {L_t^j}^{1-V_{2,t}^j}$. However, in order to establish our main result at the aggregate level, an explicit parametric specification of f is not required.

Within the above setup, we define aggregate output $\bar{Y}_t$ in period t as

$$\bar{Y}_t = \int f(K, L, V)\, dG_{t,KLV}, \tag{1.2}$$

where K, L, and V are generic random variables corresponding to capital, labor, and unobservable productivity parameters of a randomly chosen firm, respectively, and $G_{t,KLV}$ is the joint distribution of (K, L, V) across the population J_t. Thus, G_{KLV} is the explanatory variable for aggregate output. However, we do not need to model G_{KLV} but only its changes over time, since our objective is to determine the growth rate instead of the level of aggregate output.

In order to impose a structure on the evolution of the unobservable distribution of V, we introduce a set of observable firm specific attributes A_t^j with the corresponding random variable A, which are expected to be correlated with V: the age of a firm, the region or industry in which it operates, its ownership structure, and its legal form.

Further, we use A to decompose $G_{t,KLV}$ into the distributions $G_{t,V|KLA}$, $G_{t,A|KL}$, and $G_{t,KL}$. The first is the conditional distribution of V given (K, L, A), the second is the conditional distribution of A given (K, L), the third is the joint distribution of (K, L). We write

$$\begin{aligned} \bar{Y}_t &= \int\Big[\int \Big(\int f(K, L, V)\, dG_{t,V|KLA}\Big)\, dG_{t,A|KL}\Big]\, dG_{t,KL} \\ &= \int \Big(\int \bar{f}_t(K, L, A)\, dG_{t,A|KL}\Big)\, dG_{t,KL}, \end{aligned} \tag{1.3}$$

where $\bar{f}_t(K, L, A)$ is the conditional mean of output Y given (K, L, A) in period t. Thus, it is a regression function of Y on (K, L, A), which can be estimated from a cross-section of firms in period t.

From (1.3) we infer that assumptions on changes in $G_{V|KLA}$, $G_{A|KL}$, and G_{KL} are required in order to model output growth. It is easier to model the evolution of a distribution if it is symmetric, because a symmetric distribution can be well-described by its first few moments, like its mean and variance. Since the distributions of capital and labor are right-skewed in all countries, we formulate the model assumptions in terms of log capital $k_t^j := \log K_t^j$ and log labor $l_t^j := \log L_t^j$ with the corresponding random variables k and l. Further, we define $\bar{k}_t$ and $\bar{l}_t$ as the mean of k and l across J_t, respectively, and σ_t^k and σ_t^l as the corresponding standard deviations. In addition, by analogy to $G_{V|KLA}$, $G_{A|KL}$, and G_{KL}, we define $G_{V|klA}$, $G_{A|kl}$, and G_{kl}, respectively. Moreover, G_k and G_l represent marginal distributions of log capital and log labor, respectively. Finally, let $G_{\tilde{k}\tilde{l}}$ denote a component-wise standardized joint distribution of (k, l), which is defined as a joint distribution of $(\tilde{k}, \tilde{l})$, where $\tilde{k} := \frac{k-\bar{k}}{\sigma^k}$ and $\tilde{l} := \frac{l-\bar{l}}{\sigma^l}$.

In line with the aggregation approach of Hildenbrand and Kneip (2005), we impose the four following assumptions.

Assumption 1: ("Structural stability"[8] of G_{kl}) *The component-wise standardized joint distribution of log capital and log labor $G_{\tilde{k}\tilde{l}}$ is approximately equal for two consecutive periods t and $t-1$, i.e., $G_{t,\tilde{k}\tilde{l}} \approx G_{t-1,\tilde{k}\tilde{l}}$.*

It is important to note that $G_{\tilde{k}\tilde{l}}$ refers to a standardized distribution. That is, if Assumption 1 holds, the entire change in G_{kl} over two consecutive periods is fully

[8]The concept of structural stability of a distribution relies on an empirical regularity that distributions of individual variables across large populations of economic agents change very slowly over time. It has been first noticed by Pareto (1896) and introduced into macroeconomic models by Malinvaud (1993). More precisely, for a distribution of a certain parametric form, for example, the normal distribution, structural stability holds, if its normal structure prevails and its entire evolution is captured by changes in its mean and its variance. However, this concept of structural stability cannot be applied to distributions which are poorly approximated by a parametric form. In this context, Hildenbrand and Kneip (1999) proposed a nonparametric counterpart of Malinvaud's idea. Instead of keeping the parametric structure constant and allowing for changes over time in few parameters, one can keep these parameters constant and allow the shape of the distribution to vary over time. This can be achieved by simple transformations of the distribution like centering (constant mean) or standardizing (constant mean and variance). Accordingly, structural stability as defined by Hildenbrand and Kneip (1999) holds, if a centered or standardized distribution does not change over two consecutive periods.

captured by the changes in means and the variances of k_t^j and l_t^j.[9]

In order to impose the assumption on the evolution of $G_{A|kl}$ we define $k_{t,\tau}$ as the τ-quantile of the distribution $G_{t,k}$ and $l_{t,\eta}$ as the η-quantile of the distribution $G_{t,l}$.

Assumption 2: *The conditional distribution of A given $k = k_\tau$ and $l = l_\eta$ denoted by $G_{A|k_\tau l_\eta}$ is approximately equal for two consecutive periods t and $t-1$, i.e., $G_{t,A|k_\tau l_\eta} \approx G_{t-1,A|k_\tau l_\eta}$.*

Assumption 2 refers to the distribution of A across firms with log capital and log labor in the *same quantile position* (τ, η) of G_{kl} in period t and $t-1$, instead of firms with the *same values* of k and l. We employ the former specification since it increases the likelihood that we condition on the same group of firms in both periods. That is, if G_{kl} shifts over time due to a common trend, we refer to the same group of firms in both periods by conditioning on the quantile position as opposed to conditioning on the same values of k and l.

Note that one is able to verify Assumptions 1 and 2, since G_{kl} and $G_{A|kl}$ are observable in firm-level data. We document in Section 1.A below that both assumptions are supported by our data for most countries. In contrast, one is not able to falsify the following two assumptions on $G_{V|klA}$ as they concern a distribution of unobservable variables.

Let $J_t(k,l,A)$ denote the subpopulation of firms with capital k, labor l and attributes A and $\bar{V}_t(k,l,A)$ denote the mean of V across $J_t(k,l,A)$. Further, $G_{\tilde{V}|klA}$ denotes the centered distribution of V across $J_t(k,l,A)$, whereby $\tilde{V}$ corresponds to the centered variable $\tilde{V} := V - \bar{V}_t(k,l,A)$.

Assumption 3: *The distribution $G_{\tilde{V}|klA}$ is approximately equal for two periods t and $t-1$, i.e., $G_{t,\tilde{V}|klA} \approx G_{t-1,\tilde{V}|klA}$.*

---

[9]To be more precise, Hildenbrand and Kneip (2005) model the evolution of G_{kl} in terms of a distribution which is standardized by a full covariance matrix $\Sigma_t := \begin{pmatrix} (\sigma_t^k)^2 & \sigma_t^{kl} \\ \sigma_t^{kl} & (\sigma_t^l)^2 \end{pmatrix}$, instead of a component-wise standardized one, which uses the matrix $\tilde{\Sigma}_t = \begin{pmatrix} (\sigma_t^k)^2 & 0 \\ 0 & (\sigma_t^l)^2 \end{pmatrix}$. Our version of the assumption is more stringent, as it requires that the correlation between log capital and log labor is does not change significantly over two consecutive periods. The main advantage of our formulation (see Proposition and Appendix B) is the possibility to separate growth effects of changes in σ^k from growth effects of changes in σ^l.

Note that Assumption 3 is a very mild assumption since we allow for any form of heterogeneity in V across firms with different capital stocks, labor stocks, or firm characteristics. Furthermore, we even allow for heterogeneity in V across firms with the same capital stock, labor stock, and firm characteristics, as long as changes in $G_{V|klA}$ are captured by changes in $\bar{V}(k,l,A)$. In this case, we assume that $\bar{V}_t(k,l,A)$ is additively separable in (k,l) and t. More precisely,

Assumption 4: $\bar{V}_t(k,l,A)$, *can be additively factorized by* $\bar{V}_t(k,l,A) = \varphi(k,l,A) + \psi(t,A)$, *where the function* φ *is continuously differentiable in* k *and* l.

The above four assumptions allow us to derive a representation of the growth rate of the economy.

Proposition: (Hildenbrand and Kneip, 2005) *If Assumptions 1-4 hold, the growth rate of aggregate output in the economy,* $g_t := \frac{\bar{Y}_t - \bar{Y}_{t-1}}{\bar{Y}_{t-1}}$, *is given by*

$$\begin{aligned} g_t &= \beta^k_{t-1}(\log \bar{K}_t - \log \bar{K}_{t-1}) + \beta^l_{t-1}(\log \bar{L}_t - \log \bar{L}_{t-1}) && (1.4)\\ &+ \gamma^k_{t-1}\Big(\frac{\sigma^k_t - \sigma^k_{t-1}}{\sigma^k_{t-1}}\Big) + \gamma^l_{t-1}\Big(\frac{\sigma^l_t - \sigma^l_{t-1}}{\sigma^l_{t-1}}\Big) && (1.5)\\ &+ \text{(effects due to changes in } \bar{V}_{t-1}(k,l,A))\\ &+ \text{(second order terms of the Taylor expansion).} \end{aligned}$$

The coefficients β^k_{t-1}, β^l_{t-1}, γ^k_{t-1}, *and* γ^l_{t-1} *are defined in terms of partial derivatives of the regression function* $\bar{f}_{t-1}(k,l,A)$. *For* $s = \{k,l\}$ *and* $S = \{K,L\}$, β^s_{t-1}, γ^s_{t-1} *are defined by*

$$\beta^s_{t-1} = \frac{1}{\bar{Y}_{t-1}} \int \partial_s \bar{f}_{t-1}(k,l,A)\, dG_{t-1,klA}, \tag{1.6}$$

$$\gamma^s_{t-1} = \frac{1}{\bar{Y}_{t-1}} \int (s - \bar{s}_{t-1})\partial_s \bar{f}_{t-1}(k,l,A)\, dG_{t-1,klA} - \frac{\beta^s_{t-1}}{\bar{S}_{t-1}} \int (s - \bar{s}_{t-1}) \exp(s) dG_{t-1,s}, \tag{1.7}$$

and hence, do not depend on (1.1) (See, Hildenbrand and Kneip, 2005).

Remark 1: The proof of a more general result is given in Hildenbrand and Kneip (2005). The above Proposition differs from the one in Hildenbrand and Kneip (2005) in two aspects. First, our Assumption 1 relies on a component-wise standardization which makes it possible to separate growth effects of changes in σ^k from growth effects of changes in σ^l. Second, we model the aggregate relation in terms of the logarithm of aggregate variables, i.e., $\log \bar{K}$ and $\log \bar{L}$ and not the aggregates of the logarithms of individual variables, i.e., $\bar{k}$ and $\bar{l}$. This distinction yields different definitions of γ^k_{t-1} and γ^l_{t-1} and is essential to compare our model with conventional growth models, which are based on (the logarithm of) aggregate variables. See Section 1.B below for the derivations.

From the above representation we infer that the growth rate g of aggregate output does not only depend on changes in aggregate capital and aggregate labor (term (1.4)). It also depends on changes in the allocation of inputs (term (1.5)) measured by the standard deviation of log capital and log labor across firms.

The aggregate coefficients $(\beta^k_{t-1}, \gamma^k_{t-1})$ and $(\beta^l_{t-1}, \gamma^l_{t-1})$ depend on the derivatives of the regression function $\bar{f}_{t-1}$ with respect to k and l, respectively. All other variables in (1.7) are observable. The derivatives $\partial_k \bar{f}_{t-1}(k, l, A)$ and $\partial_l \bar{f}_{t-1}(k, l, A)$ can be estimated using a cross-section of firms in period $t-1$. Hence, they can be estimated independently of each other in each period. It is important to note that in the estimation of our representation of the growth rate *no time-series model* fitting takes place, which would require to include all potential growth determinants. Our estimation procedure does not require information on the growth rate of aggregate capital and labor nor the corresponding standard deviations, since the computation of aggregate coefficients is based on the estimation from a single cross-section of firms. In contrast, we are able to quantify the growth effect of changes in the distribution of inputs without specifying an exhaustive model for the aggregate growth rate. We describe the estimation methodology for these coefficients in more detail in Section 1.3.2.

Remark 2: Under Assumption 1 coefficients β^k_{t-1} and β^l_{t-1} can be interpreted as elasticities of aggregate output with respect to aggregate capital and aggregate labor, respectively. Accordingly, γ^k_{t-1} and γ^l_{t-1} are elasticities of aggregate output with respect to σ^k and σ^l, respectively. One expects β^k_{t-1} and β^l_{t-1} to be positive. However, to draw conclusions on the expected sign of γ^k_{t-1} and γ^l_{t-1} one needs to impose additional assumptions on the impact of changes in the market structure on the standard deviation of inputs. For example, if a higher degree of product market competition leads to more similarity in firm size, negative γ_{t-1}^{k} and γ_{t-1}^{l} indicate a positive relationship between growth and competition. Alternatively, we outlined above that changes in the standard deviation represent changes in the pattern of economic interactions between firms. These interactions comprise, for instance, technology spill-overs between firms. If technology diffusion is stronger among more similar firms, we expect a negative relation between spill-overs and the standard deviation of inputs and, hence, negative γ_{t-1}^{k} and γ_{t-1}^{l}.

Our theoretical result has an important implication for growth accounting. To illustrate this point, let us hypothetically claim that all variables in our model other than capital and labor do not change over time. Then, in a classical growth model, changes in $\bar{Y}$ would be in part attributed to changes in $\bar{K}$ and $\bar{L}$. However, a part of the change in $\bar{Y}$, which is not captured by the effect of changes in $\bar{K}$ and $\bar{L}$, would be attributed to changes in aggregate TFP. Such a conclusion, however, would be misleading, since we assumed that TFP did not change. From the Proposition we know that it is the effect of changes in the distribution of inputs, which is erroneously attributed to changes in TFP. Obviously, such a correct conclusion is only possible in models which allow for input heterogeneity of firms.

1.3 Empirical Analysis

In this section we structurally estimate the effects of changes in the level and allocation of capital and labor on growth separately for each of the 20 European countries in our sample.

1.3.1 Data Description

The analysis is based on European firm-level data from 2002 until 2004.[10] The data stem from the Bureau van Dijk's AMADEUS data base. It contains information from

[10] We estimate the corresponding coefficients exclusively for 2003. Yet, we need additional observations in 2002 for the Olley and Pakes (1996) estimation procedure and in 2004 for the growth accounting exercise.

firm balance sheets and covers all firms in each country. We measure output as real[11] value added. Capital and labor are measured as real fixed tangible assets and the real total cost of employees,[12] respectively. Our procedure requires that the firms have non-missing observations in 2003. Moreover, we only include countries in which data for at least 200 firms are available.

Furthermore, we include a firm's age and other variables to control for differences in economic environment across firms. In particular, we account for industry-specific and region-specific fixed effects, in that we distinguish sectors by means of two digit NACE codes and include regional dummies. Moreover, we incorporate dummy variables that capture the ownership status of a firm: (i) `quoted` takes value 1 if a firm is publicly quoted and 0 if not, while (ii) `indep1- indep9` correspond to independence indicators (defined in the AMADEUS data base) which represent different shareholder structures. Finally, we include gross investment, measured by the change in the capital stock plus depreciation, which is included as an instrument for the unobservable technology shock in the estimation procedure of Olley and Pakes (1996).

The descriptive statistics of the variables for each country in 2003 and 2004 are listed in Table 1.1. The first column indicates that the number of observations used for estimation varies substantially across countries in our sample. These differences can be attributed to different filing regulations of individual countries. For example, German companies are not legally obliged to reveal the information from their balance sheets. Hence, although the full sample for Germany covers over 800,000 firms in 2003, joint information on value added, fixed tangible assets and the number of employees is available for only roughly 6,000 German firms. In contrast, the corre-

[11]Real variables are obtained by deflating by the national output price deflators. Unfortunately, price deflators were not available at the industry level for most of the 20 European countries.

[12]We define labor in this way in order to account, to a certain extent, for differences in the quality of employees, i.e., human capital, across firms. These differences are captured by the total cost of employees, as long as firms that are characterized by the same capital stock, number of employees and the same attribute profile A, (that is, the same industry, region, age, ownership structure, etc.) but a higher human capital stock pay higher wages. We emphasize that the qualitative results do not change if we define labor as the number of employees. These results are available from authors upon request.

sponding information is available for most companies in the Spanish or Italian sample which contain about 360,000 and 117,000 observations in 2003, respectively. Analogously, means and variances of the variables differ noticeably across countries. We observe relatively large firms in Germany, the Netherlands, Austria, Great Britain and Portugal, whereas the sample covers relatively many small firms in Romania, Spain, Italy, and Sweden. Accordingly, we also observe analogue differences in the standard deviations. In all, the data reveals a substantial amount of heterogeneity both across firms within a country as well as across countries.

1.3.2 Estimation Strategy

The aggregate coefficients β_t^s and γ_t^s, $s \in \{k, l\}$ can be estimated as (suitably weighted) average derivatives in the regression of value added Y_t^j on log capital k_t^j, log labor l_t^j, and a vector of firm specific attributes A_t^j, i.e., in the model

$$Y_t^j = \bar{f}_t(k_t^j, l_t^j, A_t^j; \zeta) + u_t^j, \tag{1.8}$$

where ζ is the vector of parameters to be estimated and u_t^j is the error term with $E(u_t^j) = 0$. Hence, according to (1.6) and (1.7), once consistent estimates $\partial_s \widehat{\bar{f}_t(k, l, A}; \zeta)$ of $\partial_s \bar{f}_t(k, l, A; \zeta)$, $s \in \{k, l\}$, are obtained, one can estimate aggregate coefficients by

$$\hat{\beta}_t^k = \frac{\sum_{j \in J_t} \partial_k \widehat{\bar{f}_t(k_t^j, l_t^j}, A_t^j)}{\sum_{j \in J_t} Y_t^j}, \quad \hat{\beta}_t^l = \frac{\sum_{j \in J_t} \partial_l \widehat{\bar{f}_t(k_t^j, l_t^j}, A_t^j)}{\sum_{j \in J_t} Y_t^j}, \tag{1.9}$$

$$\hat{\gamma}_t^k = \frac{\sum_{j \in J_t} (k_t^j - \hat{\bar{k}}_t) \partial_k \widehat{\bar{f}_t(k_t^j, l_t^j, A_t^j)}}{\sum_{j \in J_t} Y_t^j} - \frac{\hat{\beta}_t^k}{\bar{K}_t} \sum_{j \in J_t} (k_t^j - \hat{\bar{k}}_t) K_t^j, \text{ and} \tag{1.10}$$

$$\hat{\gamma}_t^l = \frac{\sum_{j \in J_t} (l_t^j - \hat{\bar{l}}_t) \partial_l \widehat{\bar{f}_t(k_t^j, l_t^j, A_t^j)}}{\sum_{j \in J_t} Y_t^j} - \frac{\hat{\beta}_t^l}{\bar{L}_t} \sum_{j \in J_t} (l_t^j - \hat{\bar{l}}_t) L_t^j. \tag{1.11}$$

Our empirical strategy is focused on the model specification and estimation for $\bar{f}_t$. However, our analysis revealed that a regression of $y_t^j := \log Y_t^j$ on (k_t^j, l_t^j, A_t^j) provides a significantly better model fit and stability of results, as compared to the regression

country	n_{2003}	$\bar{Y}_{2003}$	$\bar{K}_{2003}$	$\bar{L}_{2003}$	n_{2004}	$\bar{Y}_{2004}$	$\bar{K}_{2004}$	$\bar{L}_{2004}$
Austria	1071	23.77 (112.16)	27.06 (127.48)	18.10 (62.71)	1364	23.88 (113.94)	27.30 (130.30)	14.04 (43.78)
Belgium	10980	12.16 (123.64)	6.71 (32.04)	4.65 (13.92)	11036	12.15 (123.74)	6.72 (31.40)	5.16 (15.68)
Bosnia & H.	2573	0.42 (2.81)	1.36 (9.77)	0.12 (0.35)	2862	0.40 (2.64)	1.22 (7.59)	0.13 (0.38)
Bulgaria	5818	0.31 (2.73)	0.76 (4.08)	0.16 (0.59)	5955	0.31 (2.74)	0.78 (4.03)	0.18 (0.66)
Czech R.	11494	1.26 (14.42)	2.00 (9.66)	0.62 (1.61)	15799	1.27 (13.46)	2.00 (9.83)	0.62 (1.67)
Denmark	20426	2.92 (78.92)	1.36 (7.84)	1.17 (4.93)	21782	2.98 (77.82)	1.37 (7.80)	1.18 (4.78)
Estonia	7666	0.23 (1.80)	0.24 (0.97)	0.10 (0.24)	8083	0.24 (1.81)	0.26 (1.06)	0.11 (0.30)
Finland	32401	1.70 (39.11)	0.80 (5.32)	0.67 (2.85)	30328	1.70 (39.32)	0.73 (4.81)	0.79 (3.22)
France	157141	1.91 (49.72)	0.74 (4.91)	1.15 (4.08)	168079	2.05 (52.41)	0.73 (4.79)	1.21 (4.35)
Germany	6076	71.49 (840.30)	61.75 (273.70)	45.14 (188.75)	7623	68.27 (778.79)	62.81 (296.03)	34.94 (130.98)
Great Britain	41649	18.93 (263.41)	13.44 (88.36)	8.67 (34.79)	37666	19.16 (269.78)	14.75 (98.52)	11.58 (46.10)
Italy	117111	2.39 (86.51)	1.46 (6.62)	1.06 (3.62)	75392	1.98 (24.33)	1.56 (7.38)	1.98 (6.54)
Netherlands	7365	24.51 (329.13)	19.99 (109.56)	16.70 (71.05)	7375	25.34 (347.45)	20.30 (115.71)	17.98 (78.17)
Norway	12051	1.42 (45.92)	1.79 (9.65)	0.54 (1.30)	14299	1.43 (46.21)	1.75 (9.62)	0.68 (1.98)
Poland	10571	2.61 (26.13)	3.34 (13.32)	0.92 (2.12)	11188	2.55 (25.34)	3.82 (14.85)	1.10 (2.54)
Portugal	1451	9.96 (84.90)	19.79 (147.20)	6.11 (25.62)	1487	9.33 (84.48)	21.40 (154.91)	6.00 (25.17)
Romania	49018	0.10 (2.35)	0.10 (0.45)	0.05 (0.16)	66230	0.10 (2.40)	0.12 (0.51)	0.04 (0.14)
Slovakia	2042	1.63 (11.35)	4.16 (30.28)	0.84 (2.30)	2557	2.41 (22.98)	3.23 (23.58)	0.83 (2.49)
Spain	357410	0.96 (34.63)	0.49 (2.25)	0.31 (1.07)	360517	1.00 (37.23)	0.52 (2.37)	0.34 (1.18)
Sweden	123058	1.56 (42.47)	0.73 (5.52)	0.40 (1.78)	125725	1.47 (38.70)	0.74 (5.55)	0.44 (1.91)

All values are stated in millions of EUR.

Table 1.1: Descriptive statistics of the AMADEUS data for 20 European countries.

of Y_t^j on (k_t^j, l_t^j, A_t^j). Consequently, we estimate derivatives of $\bar{f}_t$ from the model

$$y_t^j = \bar{h}_t(k_t^j, l_t^j, A_t^j; \theta) + \varepsilon_t^j, \tag{1.12}$$

where θ is the vector of parameters to be estimated and ε_t^j is the error term with $E(\varepsilon_t^j) = 0$. In doing so, we use the fact that $\partial_s \bar{f}_t(k_t^j, l_t^j, A_t^j; \hat{\zeta}) = Y_t^j \partial_s \bar{h}_t(k_t^j, l_t^j, A_t^j; \hat{\theta})$, if $\hat{\zeta}$ and $\hat{\theta}$ are consistent estimates of ζ and θ, respectively. Our basic specification for $\bar{h}_t$ is linear in (k, l, A) and can be estimated using OLS. Further, we analyze the robustness of our results in two ways. First, we control for possible simultaneity between ε_t^j and (k, l) using the Olley and Pakes (1996) method. Second, we extend our analysis to a partially linear specification of $\bar{h}_t$, in which the relationship between y and (k, l) is modeled nonparametrically. Doing this, we avoid a parametric misspecification of $\bar{h}_t$.

The loglinear model

Our basic specification for $\bar{h}_t$ is the loglinear model

$$y_t^j = \theta_0 + \theta^k k_t^j + \theta^l l_t^j + \theta_A' A_t^j + \varepsilon_t^j, \tag{1.13}$$

which implies that $\partial_k \widehat{\bar{f}_t(k_t^j, l_t^j}, A_t^j) = \hat{\theta}^k Y_t^j$ and $\partial_l \widehat{\bar{f}_t(k_t^j, l_t^j}, A_t^j) = \hat{\theta}^l Y_t^j$.[13] These quantities are then imputed into (1.9) - (1.11), in order to calculate aggregate parameters.

In the simplest case, (1.13) can be estimated by the OLS method from a single cross-section in 2003. However, the vast literature on estimation of production functions from plant-level data points out that OLS may suffer from a simultaneity problem. This problem arises if there is a contemporaneous correlation between the demand for inputs k_t^j, l_t^j and the realization of the unobservable technology shock contained in ε_t^j. In such a case, estimates $\hat{\theta}^k$ and $\hat{\theta}^l$, and, hence, $\hat{\beta}^k$ and $\hat{\beta}^l$ would be biased. There are several approaches to correct for simultaneity between (k_t^j, l_t^j) and ε_t^j and all of them put additional restrictions on the data. For instance, Olley and Pakes (1996) propose a method, which uses changes in firm's investment decision as a proxy for the productivity shock. However, only firms with non-missing data for 2002 and 2003 on

[13]Note that in this model, $\hat{\beta}^k = \hat{\theta}^k$ and $\hat{\beta}^l = \hat{\theta}^l$.

value added, capital, labor, and investment can be used for estimation. Depending on the country, this requirement involves an elimination of up to 70% of the companies from our sample of firms with non missing data on value added, capital, and labor in 2003. Moreover, the above method may introduce a sample selection bias, if dropping out of the sample between 2002 and 2003 is non-random. Following the same idea, Levinsohn and Petrin (2003) suggest the use of intermediate inputs instead of the investment variable as a proxy.[14] Finally, as described in Blundell and Bond (2000), the simultaneity problem in estimation of production function can also be bypassed by a GMM system estimator, though it requires a long time-series of cross-sections and is therefore not attractive for our analysis.

Being aware of problems mentioned above, we consistently estimate (1.13) following Olley and Pakes (1996) in controlling for both simultaneity bias and sample attrition. The method is based on a two-step procedure and requires following assumptions: (i) labor is the only input which contemporaneously responds to a technology shock, (ii) capital stock is predetermined and hence uncorrelated with a contemporary technology shock, (iii) changes in corporate investment decisions depend on the contemporaneous technology shock, the age and the capital stock of a firm, (iv) investments are monotonically increasing in the technology shock for a given value of age and capital. Under these assumptions, the technology shock can be instrumented as a function of capital, age, and investment. The estimation of this function is carried out by a series estimator.

Semiparametric model

In order avoid a misspecification of the relationship between y and (k, l, A) we model $\bar{h}_t$ semiparametrically and include an interaction term

$$y_t^j = \theta_0 + \bar{h}_t^k(k_t^j) + \bar{h}_t^l(l_t^j) + \theta^{kl} k_t^j l_t^j + \theta'_A A_t^j + \varepsilon_t^j, \tag{1.14}$$

[14]They motivate their choice by weaker data requirements and argue that an adjustment in intermediate inputs is likely to have better properties as an instrument for a technology shock than an adjustment in investment. Interestingly, the approach of Levinsohn and Petrin (2003) requires even more firms to be eliminated from our sample due to the very large number of firms with missing data on the use of materials.

where $\bar{h}_t^k$ and $\bar{h}_t^l$ are differentiable in k and l, respectively. We model $\bar{h}_t^k$ as a quadratic splines function with D^k knots $d_1^k < d_2^k < \cdots < d_{D^k}^k$. Defining basis functions $b_i^k(k) = \max\{0, k - d_i^k\}^2$, we obtain $\bar{h}_t^k(k) = \theta_1^k k + \theta_2^k k^2 + \sum_{i=1}^{D^k} \theta_{3,i}^k b_i^k(k)$. Analogously, we model $\bar{h}_t^l$ as $\bar{h}_t^l(l) = \theta_1^l l + \theta_2^l l^2 + \sum_{i=1}^{D^l} \theta_{3,i}^l b_i^l(l)$. All coefficients in (1.14) can be estimated by the OLS method. Accordingly, $\partial_k \bar{f}_t(k_t^j, l_t^j, A_t^j)$ can be estimated as

$$\partial_k \widehat{\bar{f}_t(k_t^j, l_t^j, A_t^j)} = \Big(\hat{\theta}_1^k + 2\hat{\theta}_2^k k_t^j + \hat{\theta}^{kl} l_t^j + 2\sum_{i=1}^{D^k} \hat{\theta}_{3,i}^k \max\{0, k_t^j - d_i^k\}\Big) Y_t^j.$$

Similarly, one obtains

$$\partial_l \widehat{\bar{f}_t(k_t^j, l_t^j, A_t^j)} = \Big(\hat{\theta}_1^l + 2\hat{\theta}_2^l l_t^j + \hat{\theta}^{kl} k_t^j + 2\sum_{i=1}^{D^l} \hat{\theta}_{3,i}^l \max\{0, l_t^j - d_i^l\}\Big) Y_t^j.$$

The optimal number of knots and their position is obtained by the minimization of the Mallows' C_p criterion (see Mallows, 1973) using the knot deletion method as described by Fan and Gijbels (1996, p. 42).[15]

Statistical significance of the aggregate coefficients

Confidence intervals for the aggregate coefficients as well as standard errors of the estimates are determined by bootstrap. For i.i.d. bootstrap resamples $(Y_t^{j*}, k_t^{j*}, l_t^{j*}, A_t^{j*})$, the distribution of $(\hat{\beta}_t^k - \beta_t^k)$ is approximated by the conditional distribution of $(\hat{\beta}_t^{k*} - \hat{\beta}_t^k)$ given $(Y_t^j, k_t^j, l_t^j, A_t^j)$, where $\hat{\beta}_t^{k*}$ is the estimate of β_t^k based on the bootstrap sample. We asses the significance of β_t^k on the basis of the 95% confidence interval, $[\hat{\beta}_t^k - q_{0.975}^*, \hat{\beta}_t^k - q_{0.025}^*]$, where q_α^* is the α-quantile of the distribution of $(\hat{\beta}_t^{k*} - \hat{\beta}_t^k)$. Analogously, we compute confidence intervals for β_t^l, γ_t^k, and γ_t^l. Distributional effects

[15]Knot deletion is an iterative procedure. We start with a large number $\bar{D}^k$ of initial knots for k, i.e., $d_1^k < d_2^k < \cdots < d_{\bar{D}^k}^k$, which divide the domain of k into intervals $[d_i^k, d_{i+1}^k]$ with approximately equal number of observations. Similarly, we determine the corresponding $\bar{D}^l$ initial knots for l. In step 0, we estimate (1.14) by the OLS method and obtain $\bar{D} = \bar{D}^k + \bar{D}^l$ estimated spline coefficients $\hat{\theta}_{3,1}^k, \ldots, \hat{\theta}_{3,\bar{D}^k}^k, \hat{\theta}_{3,1}^l, \ldots, \hat{\theta}_{3,\bar{D}^l}^l$ with the corresponding t-values, $t := \hat{\theta}/SE(\hat{\theta})$. At step 1, we delete the knot with the lowest absolute t-value at step 0 and reestimate (1.14) using $\bar{D} - 1$ knots. We repeat this process $\bar{D}$ times until no knots are left. At each step r, $0 \leqslant r \leqslant \bar{D}$, we compute the *residual sum of squares* $RRS_r = \sum_{j=1}^n (\hat{\varepsilon}_t^j)^2$. Finally, we choose the model with the lowest value for Mallows' C_p defined by $C_r := RSS_r + 3(\bar{D} + 6 + n_A - r)\hat{\sigma}_0^2$, where n_A is the number of attributes in A_t^j and $\hat{\sigma}_0$ is the estimated standard deviation of ε_t^j at the 0^{th} model.

are statistically significant, if the confidence interval for γ_t^k or γ_t^l does not include zero. The consistency proof of such a naive bootstrap in the context of average derivative estimation can be found in Härdle and Hart (1992).

1.3.3 Empirical Results

In the following we present the results for the estimation of β^k, β^l, γ^k, and γ^l. We report results based on the OLS estimation of (1.13) in Table 1.2. The first two columns of the table reveal that, as expected, changes in the levels of aggregate capital and labor have a positive significant effect on growth in all countries. Further, the capital coefficient appears to be higher for transition than for developed countries. Overall, the estimated aggregate output elasticities with respect to aggregate capital and labor, i.e., $\hat{\beta}^k$ and $\hat{\beta}^l$, are comparable with those obtained by other studies.[16] More interestingly, we find that distributional effects of capital or labor, associated with γ^k and γ^l, are significant at 1% level in all countries. These coefficients are displayed in the last two columns of Table 2. Further, the distributional effects of capital are negative and higher (in absolute value) than the corresponding level effects associated with β^k. As for distributional effects of labor, they turn out to be negative and significant at 1% level for all countries except from Austria, Czech Republic, Portugal and Slovakia. For Portugal they are positive and significant at the 5% level. Summing up, distributional effects of capital and labor, which have been overlooked in the growth literature so far, are statistically and economically significant.

We investigate the robustness of this finding by controlling for potential simultaneity and misspecification of the functional form. Table 1.3 reports the estimation results according to the Olley and Pakes (1996) method. Overall, the estimates are similar to the OLS estimates but exhibit higher standard errors. We infer that the simultaneity problem is of less importance in our sample. In particular, γ^k is still negative and significant for all countries. Moreover, apart from Germany and Roma-

[16] Recall that under this specification $\hat{\beta}^k = \hat{\theta}^k$ and $\hat{\beta}^l = \hat{\theta}^l$. Hence, we can compare our estimates with those obtained in studies on production function estimation from the firm-level data, e.g., Olley and Pakes (1996), Levinsohn and Petrin (2003), and Blundell and Bond (2000).

country	$\hat{\beta}^k$	$\hat{\beta}^l$	$\hat{\gamma}^k$	$\hat{\gamma}^l$
Austria	0.151 (0.016)	0.788 (0.025)	-0.190 (0.034)*	-0.037 (0.054)
Belgium	0.140 (0.006)	0.749 (0.008)	-0.293 (0.020)*	-0.250 (0.030)*
Bosnia & H.	0.212 (0.011)	0.581 (0.015)	-0.351 (0.039)*	-0.166 (0.036)*
Bulgaria	0.234 (0.009)	0.639 (0.010)	-0.268 (0.027)*	-0.190 (0.063)*
Czech R.	0.140 (0.004)	0.811 (0.007)	-0.183 (0.011)*	0.035 (0.026)
Denmark	0.116 (0.004)	0.747 (0.006)	-0.181 (0.012)*	-0.149 (0.024)*
Estonia	0.185 (0.008)	0.715 (0.009)	-0.278 (0.019)*	-0.210 (0.029)*
Finland	0.147 (0.002)	0.778 (0.003)	-0.299 (0.014)*	-0.090 (0.011)*
France	0.111 (0.001)	0.854 (0.002)	-0.232 (0.005)*	-0.038 (0.007)*
Germany	0.136 (0.007)	0.803 (0.011)	-0.130 (0.017)*	-0.107 (0.037)*
Great Britain	0.132 (0.003)	0.783 (0.004)	-0.248 (0.010)*	-0.057 (0.016)*
Italy	0.131 (0.002)	0.732 (0.002)	-0.179 (0.004)*	-0.058 (0.007)*
Netherlands	0.119 (0.007)	0.832 (0.010)	-0.171 (0.017)*	-0.158 (0.035)*
Norway	0.091 (0.003)	0.804 (0.006)	-0.210 (0.011)*	-0.123 (0.018)*
Poland	0.152 (0.006)	0.774 (0.009)	-0.213 (0.012)*	-0.077 (0.021)*
Portugal	0.130 (0.017)	0.818 (0.022)	-0.170 (0.032)*	0.132 (0.060)*
Romania	0.252 (0.003)	0.667 (0.004)	-0.241 (0.008)*	-0.319 (0.010)*
Slovakia	0.156 (0.013)	0.743 (0.020)	-0.193 (0.037)*	0.136 (0.086)
Spain	0.115 (0.001)	0.841 (0.001)	-0.181 (0.003)*	-0.103 (0.006)*
Sweden	0.148 (0.001)	0.766 (0.002)	-0.351 (0.008)*	-0.089 (0.012)*

Bootstrapped standard errors are given in parentheses. Asterisks denote statistical significance of distributional effects at the 5% level.

Table 1.2: Estimated values of aggregate coefficients based on OLS production function estimation.

nia, the distributional effects of capital are again stronger (in absolute value) than the corresponding level effect. The distributional effects of labor are negative and significant in 13 out of 20 countries. The results for the semiparametric estimation are reported in Table 1.4. We observe that the estimates of β^k exceed the corresponding OLS estimates in most countries. In contrast, $\hat{\beta}^l$ are comparable to the OLS counterparts. At least one of the distributional effects, i.e., γ^k or γ^l, is significant in all countries apart from the Czech Republic and Slovakia. Interestingly, accounting for a more flexible functional form yields positive significant distributional effect of capital in Denmark, Italy and Norway. In contrast, γ^k is negative significant for eleven countries. Besides, the distributional effects of capital are smaller than the ones resulting from the loglinear model. As opposed to previous models, they are also lower than the corresponding level effects. As for distributional effects of labor, they are significantly negative in ten countries and significantly positive in Portugal. Summing up, the importance of the distributional effects, which are the main focus of this chapter, is robust to simultaneity and parametric misspecification.

country	$\hat{\beta}^k$	$\hat{\beta}^l$	$\hat{\gamma}^k$	$\hat{\gamma}^l$
Austria	0.165 (0.067)	0.795 (0.087)	-0.240 (0.127)*	-0.010 (0.062)
Belgium	0.159 (0.029)	0.715 (0.009)	-0.298 (0.057)*	-0.184 (0.037)*
Bosnia & H.	0.266 (0.076)	0.509 (0.020)	-0.195 (0.86)*	-0.260 (0.068)*
Bulgaria	0.286 (0.042)	0.560 (0.017)	-0.304 (0.062)*	-0.089 (0.072)
Czech R.	0.111 (0.045)	0.752 (0.014)	-0.124 (0.051)*	0.029 (0.040)
Denmark	0.121 (0.039)	0.760 (0.008)	-0.166 (0.053)*	-0.095 (0.017)*
Estonia	0.185 (0.020)	0.685 (0.012)	-0.209 (0.025)*	-0.080 (0.034)*
Finland	0.156 (0.017)	0.763 (0.005)	-0.282 (0.035)*	-0.067 (0.013)*
France	0.119 (0.009)	0.829 (0.003)	-0.228 (0.018)*	-0.031 (0.008)*
Germany	0.117 (0.038)	0.744 (0.016)	-0.081 (0.035)*	-0.020 (0.044)
Great Britain	0.155 (0.035)	0.782 (0.005)	-0.285 (0.067)*	-0.038 (0.019)*
Italy	0.163 (0.017)	0.705 (0.003)	-0.173 (0.018)*	-0.061 (0.007)*
Netherlands	0.180 (0.031)	0.758 (0.013)	-0.213 (0.041)*	-0.051 (0.034)
Norway	0.064 (0.007)	0.835 (0.008)	-0.109 (0.012)*	-0.059 (0.006)*
Poland	0.123 (0.046)	0.741 (0.011)	-0.164 (0.065)*	-0.091 (0.032)*
Portugal	0.126 (0.051)	0.832 (0.041)	-0.236 (0.101)*	0.007 (0.062)
Romania	0.147 (0.044)	0.629 (0.006)	-0.101 (0.030)*	-0.252 (0.014)*
Slovakia	0.158 (0.053)	0.682 (0.028)	-0.186 (0.072)*	0.234 (0.135)
Spain	0.121 (0.010)	0.817 (0.002)	-0.173 (0.015)*	-0.063 (0.007)*
Sweden	0.154 (0.007)	0.759 (0.002)	-0.353 (0.018)*	-0.070 (0.012)*

Bootstrapped standard errors are given in parentheses. Asterisks denote statistical significance of distributional effects at the 5% level.

Table 1.3: Estimated values of aggregate coefficients based on the Olley and Pakes (1996) method.

The negative impact of changes in the standard deviation of inputs in most countries supports the intuition outlined in Remark 2. First, under the assumption that a higher degree of product market competition among firms is associated with more similarity in firm size, i.e., smaller standard deviations of capital and labor, we find a positive relationship between competition and economic growth. This positive relation

country	$\hat{\beta}^k$	$\hat{\beta}^l$	$\hat{\gamma}^k$	$\hat{\gamma}^l$
Austria	0.171 (0.030)	0.779 (0.035)	-0.095 (0.045)*	-0.212 (0.061)*
Belgium	0.142 (0.011)	0.813 (0.014)	-0.097 (0.018)*	-0.231 (0.041)*
Bosnia & H.	0.240 (0.047)	0.729 (0.040)	-0.340 (0.057)*	0.109 (0.077)
Bulgaria	0.295 (0.036)	0.725 (0.041)	-0.095 (0.053)*	-0.050 (0.087)
Czech R.	0.257 (0.025)	0.793 (0.020)	-0.024 (0.039)	0.067 (0.038)
Denmark	0.174 (0.015)	0.796 (0.013)	0.038 (0.022)*	-0.220 (0.034)*
Estonia	0.187 (0.016)	0.775 (0.020)	-0.119 (0.025)*	-0.109 (0.043)
Finland	0.160 (0.010)	0.833 (0.010)	-0.095 (0.017)*	-0.090 (0.021)*
France	0.119 (0.003)	0.870 (0.004)	-0.059 (0.006)*	-0.024 (0.011)*
Germany	0.178 (0.013)	0.815 (0.016)	-0.006 (0.020)	-0.100 (0.044)*
Great Britain	0.211 (0.008)	0.797 (0.009)	-0.066 (0.012)*	-0.125 (0.021)*
Italy	0.153 (0.007)	0.820 (0.006)	-0.027 (0.021)	-0.063 (0.013)*
Netherlands	0.170 (0.019)	0.829 (0.022)	-0.002 (0.038)	-0.115 (0.050)*
Norway	0.141 (0.010)	0.856 (0.011)	0.060 (0.016)*	-0.050 (0.027)
Poland	0.156 (0.017)	0.856 (0.017)	-0.130 (0.031)*	-0.024 (0.033)
Portugal	0.231 (0.058)	0.805 (0.074)	-0.045 (0.037)	0.149 (0.084)*
Romania	0.209 (0.009)	0.693 (0.008)	-0.264 (0.018)*	-0.206 (0.014)*
Slovakia	0.309 (0.060)	0.730 (0.053)	-0.082 (0.089)	0.141 (0.103)
Spain	0.164 (0.004)	0.831 (0.003)	-0.001 (0.006)	-0.142 (0.009)*
Sweden	0.173 (0.004)	0.820 (0.005)	-0.095 (0.008)*	-0.047 (0.014)*

Bootstrapped standard errors are given in parentheses. Asterisks denote statistical significance of distributional effects at the 5% level.

Table 1.4: Estimated values of aggregate coefficients based on the semiparametric specification.

is also found in the literature, for instance, by Nicoletti and Scarpetta (2003).

Second, changes in the distribution of inputs capture changes in the pattern of economic interactions between firms. In particular, the literature on economic growth emphasizes the importance of technology spill-overs among firms in developed economies. A standard assumption in the literature is that technology spill-overs are more likely between firms that are more similar in terms of the inputs they use in the production process.[17] Accordingly, an increase in the standard deviation of capital or labor corresponds to less intensive technology spill-overs and, hence, to lower growth rates.

[17]Theoretical models by Basu and Weil (1998) and Acemoglu and Zilibotti (2000) show that international technology diffusion is stronger if firms employ more similar capital-labor ratios in production. An empirical evidence in favor of this result is provided by Keller (2004).

1.4 Growth Accounting

We exploit the economic significance of the distributional effects outlined above to refine conventional growth accounting exercises. That is, we explore whether cross-country growth differences can be explained by differences in changes in the allocation of capital and labor. Their explanatory power depends on the cross-country heterogeneity in γ^k and γ^l as well as in the growth rates of the standard deviations of the inputs.

To measure the success of a model in explaining cross-country growth differences we follow the tradition of variance decomposition. That is, analog to Caselli (2005), we compute the explanatory power of the changes in the aggregate input levels as

$$\mathtt{S1} = \frac{var(\hat{g}_{1,t})}{var(g_t)} \tag{1.15}$$

where

$$\hat{g}_{1,t} = \hat{\beta}^k_{t-1}(\log \bar{K}_t - \log \bar{K}_{t-1}) + \hat{\beta}^l_{t-1}(\log \bar{L}_t - \log \bar{L}_{t-1}).$$

The residual of this indicator, $1 - \mathtt{S1}$, is the explanatory power of changes in TFP. However, we know from the Proposition that part of the residual changes should not be associated to changes in the production technology (TFP), but instead, to changes in the higher moments of the distribution of capital and labor across firms. Accordingly, our approach which takes firm-level heterogeneity in the inputs into account leads to a different growth accounting model:

$$\mathtt{S2} = \frac{var(\hat{g}_{2,t})}{var(g_t)}, \tag{1.16}$$

where

$$\hat{g}_{2,t} = \hat{\beta}^k_{t-1}(\log \bar{K}_t - \log \bar{K}_{t-1}) + \hat{\beta}^l_{t-1}(\log \bar{L}_t - \log \bar{L}_{t-1}) + \hat{\gamma}^k_{t-1}\Big(\frac{\sigma^k_t - \sigma^k_{t-1}}{\sigma^k_{t-1}}\Big) + \hat{\gamma}^l_{t-1}\Big(\frac{\sigma^l_t - \sigma^l_{t-1}}{\sigma^l_{t-1}}\Big).$$

In addition to the estimated aggregate coefficients growth accounting requires data on the growth rate of aggregate output, aggregate capital, aggregate labor and the standard deviations of log capital and log labor. Since the estimation of coefficients relies on data in 2003 (corresponding to $t-1$) we focus on growth rates from 2003 to 2004. All of the required information is available in the AMADEUS data base. However, the computation of aggregate output and inputs from the cross-section of firms yields implausibly high growth rates of these variables (see Table 1.1). Therefore, we employ information on growth rates of aggregate variables from the standard cross-country data sets. In particular, we employ Penn World Tables and follow Caselli (2005) in measuring output as real GDP per capita in PPP and computing the aggregate capital stock from the corresponding investment series using the perpetual inventory method and by assuming yearly depreciation rate of 6%. Since aggregate labor in 2004 is not available in Penn World Tables, we measure aggregate labor as total number of employees from the Eurostat data base. Obviously, the information on the standard deviations of log capital and log labor has to be obtained from the firm-level data base. Unfortunately, required aggregate data for Bosnia and Herzegovina are not available and we are forced to omit this country in our analysis. The growth rates of the variables employed in the growth accounting exercise are reported in Table 1.5.

We derive S1 and S2 based on the three different estimators outlined in the last section. In particular, we find that the aggregate capital and labor explain 28% of the cross-country growth differences based on the OLS estimates ($\texttt{S1}_{OLS} = 0.28$), 29% based on the Olley and Pakes (1996) method ($\texttt{S1}_{OP} = 0.29$), and 40% based on the semiparametric model ($\texttt{S1}_{SP} = 0.40$). These results are consistent with the corresponding findings in the conventional growth accounting literature. If we additionally take the distributional effects into consideration, we are able to explain an additional 17%, 13%, and 6% of the growth differences across countries, respectively ($\texttt{S2}_{OLS} = 0.45$, $\texttt{S2}_{OP} = 0.42$, $\texttt{S2}_{SP} = 0.46$). Recall that, our aggregate coefficients are not estimated by fitting changes in aggregate levels and standard deviations to output growth rates, but are computed from a structural estimation based on firm-level data. Hence, in contrast to standard goodness-of-fit measures, the explanatory power could drop if we additionally account for distributional effects. This would be the

country	g_{04}	$\log \frac{\bar{K}_{04}}{\bar{K}_{03}}$	$\log \frac{\bar{L}_{04}}{\bar{L}_{03}}$	$\frac{\sigma^k_{04}-\sigma^k_{03}}{\sigma^k_{03}}$	$\frac{\sigma^l_{04}-\sigma^l_{03}}{\sigma^l_{03}}$
Austria	2.14	-1.31	0.57	-2.46	-1.89
Belgium	2.46	3.52	0.65	0.61	-0.76
Bosnia & H.	-	-	-	-5.14	-6.20
Bulgaria	5.02	10.02	2.59	-0.62	-1.38
Czech R.	3.10	4.73	-0.28	-0.43	2.33
Denmark	1.71	2.22	0.00	0.79	-1.00
Estonia	7.73	-0.54	0.25	1.24	0.48
Finland	3.47	2.75	0.41	-3.33	-0.22
France	1.97	5.03	0.05	0.46	0.38
Germany	1.66	1.13	0.42	1.22	0.27
Great Britain	2.75	1.93	1.00	1.52	0.56
Italy	1.09	0.28	0.37	3.78	10.14
Netherlands	1.23	2.25	-1.42	-0.21	1.79
Norway	2.20	9.26	0.47	0.83	1.39
Poland	5.31	6.36	1.31	-0.28	0.66
Portugal	0.38	1.26	0.09	0.22	3.50
Romania	8.68	1.64	0.39	-5.42	1.86
Slovakia	3.50	9.25	0.27	-10.04	-4.29
Spain	1.61	1.95	3.42	0.06	-0.92
Sweden	3.58	-1.27	-0.57	1.61	1.04

Table 1.5: Growth rates in 2004 (in %) used in the growth accounting exercise.

case if the changes in σ^k and σ^l were negatively correlated with omitted factors that explain GDP-growth. Consequently, distributional effects of capital and labor across firms help explain a significant part of variation in growth across the 19 European countries.

We analyze the robustness of the above result in two different ways. First, we redo the growth accounting exercise by excluding one country at a time. We repeat this procedure for all countries. Doing this, we obtain very similar results as the ones from the unrestricted sample. Second, we extend the sample period to 2002-2004, which virtually does not change our results. In all, the growth accounting results are robust to variations in the cross-section as well as in the time-series dimension.

Overall, we conclude that accounting for distributional effects of capital and labor helps explain an additional 6-17% of the cross-country variation in output growth among the 19 European countries. Thus, a growth accounting model which is based on the correct treatment of firm heterogeneity improves the explanatory power of the production inputs and reduces the relevance of the residual TFP measure.

1.5 Summary and Conclusions

In this chapter, we propose a growth model to examine the effect of distributional changes of capital and labor on economic growth. We show that the growth rate of an economy depends not only on changes in the aggregate level of capital and labor, but also on changes in the allocation of these inputs across firms, which we measure by standard deviations of capital and labor. Our empirical analysis, based on European firm-level data, reveals that changes in the allocation of capital and labor have pronounced effects on GDP-growth in almost all of the 20 European countries. This striking result revises the rather unimportant role of capital and labor distributions in explaining income and growth differences across countries as documented, for instance, by Caselli (2005). Moreover, it suggests that conventional TFP measures misleadingly capture growth effects stemming from changes in the standard deviations of capital and labor. In fact, our framework allows us to assess the explanatory power of higher moments of the input distributions and, therefore, reassess the explanatory power of TFP. In this regard, we refine conventional growth accounting exercises by controlling for cross-country differences in aggregate input levels *and* input allocations.

Our empirical results reveal that distributional effects from firm-level heterogeneity in the inputs are statistically and economically significant in almost all countries. In particular, we find that higher standard deviations in labor and capital have negative effects on output growth. This finding is consistent with a positive relationship between competition and growth if more competition is associated with more similarity in firm size and, hence, lower standard deviations in capital and labor among firms. Our findings are also consistent with the fact that if firms are getting similar, the technology spill-overs are more intensive, which promotes economic growth.

Finally, in a growth accounting exercises we show that distributional effects of capital and labor help explain an additional 6-17% of cross-country growth differences among the 19 European countries.

1.A Appendix: Empirical Analysis of Assumptions 1 and 2

Assumption 1

We aim to analyze whether the standardized joint distribution of log capital and log labor, i.e., $G_{\tilde{k}l}$, changes sufficiently slowly over time, so that it can be regarded as approximately equal for 2003 and 2004. In order to answer this question, we apply a nonparametric kernel-based test of closeness between two distribution functions as proposed by Li (1996). Under the null hypothesis that two distributions are equal, the test statistic T, which relies on the integrated squared difference between $G_{2003,\tilde{k}l}$ and $G_{2004,\tilde{k}l}$, has a standard normal distribution. However, the asymptotic distribution of T under the null hypothesis has a slow rate of convergence to the the standard normal distribution. In order to account for this finite sample bias, we perform the bootstrap procedure to approximate the distribution of T. We repeat the following procedure $B = 500$ times: Out of the pooled sample $\{(k^1_{2003}, l^1_{2003}), \ldots, (k^{n_{2003}}_{2003}, l^{n_{2003}}_{2003}); (k^1_{2004}, l^1_{2004}), \ldots, (k^{n_{2004}}_{2004}, l^{n_{2004}}_{2004})\}$ two samples $\{(k^{*1}, l^{*1}), \ldots, (k^{*n_{2003}}, l^{*n_{2003}})$ and $\{(k^{*1}, l^{*1}), \ldots, (k^{*n_{2004}}, l^{*n_{2004}})$ are randomly drawn with replacement. Then, based on the new samples the test statistic T^*_b is computed. The empirical distribution of T under the null hypothesis is then estimated from the sample $\{T^*_1, \ldots, T^*_B\}$. The consistency of the bootstrap in this context is proven by Li et al (2007). Moreover, bandwidth parameters used for testing were obtained through the Sheather and Jones (1991) method.

Assumption 1 is well supported by the Amadeus data. The test results for 20 countries are given in Table 1.6. They indicate that changes in $G_{\tilde{k}l}$ from 2003 to 2004 can be indeed regarded as statistically insignificant for 17 out of 20 countries in our sample. We reject equality of $G_{2003,\tilde{k}l}$ and $G_{2004,\tilde{k}l}$ only for Finland, Italy, and Romania.

country	test stat.	emp. p-value	as. p-value
Austria	-1.741	0.950	0.959
Belgium	-0.454	0.591	0.675
Bosnia & H.	-2.069	0.976	0.981
Bulgaria	0.047	0.456	0.481
Czech R.	-1.659	0.922	0.951
Denmark	0.259	0.310	0.398
Estonia	-1.231	0.856	0.891
Finland	3.973*	0.001	0.000
France	-0.193	0.502	0.577
Germany	1.343	0.057	0.090
Great Britain	1.512	0.077	0.065
Italy	12.522*	0.000	0.000
Netherlands	-1.966	0.951	0.975
Norway	-0.565	0.696	0.714
Poland	-1.970	0.975	0.976
Portugal	-1.889	0.970	0.971
Romania	3.161*	0.013	0.001
Slovakia	-0.892	0.733	0.814
Spain	-1.067	0.823	0.857
Sweden	1.562	0.072	0.059

Asterisks denote that changes in the (coordinate-wise) standardized joint distribution of log capital and log labor from 2003 to 2004 were statistically significant at the 5% level.

Table 1.6: Empirical verification of Assumption 1 using the Li (1996) test for equality of distributions.

Assumption 2

Recall that we denote by $k_{t,\tau}$ the τ-quantile of the distribution $G_{t,k}$ and by $l_{t,\eta}$ the η-quantile of the distribution $G_{t,l}$. We analyze whether for all $0 < \tau < 1$ and $0 < \eta < 1$ the conditional distribution of attributes given $k = k_\tau$ and $l = l_\eta$, i.e., $G_{A|k_\tau,l_\eta}$ changed significantly from 2003 to 2004.

In our analysis A_t^j contains company age, industry and regional dummies, independence indicators and a dummy for being publicly quoted. Since among these variables solely the age of a company age is a continuous variable, while verifying Hypothesis 2, we concentrate on the evolution of the conditional distribution of age, i.e., $G_{age|k_\tau l_\eta}$. We study the evolution this distribution for $(\tau, \eta) \in \{(0.1, 0.1), (0.25, 0.25), (0.5, 0.5), (0.75, 0.75), (0.9, 0.9)\}$.[18] In order to assess the significance of changes in $G_{age|k_\tau l_\eta}$ from 2003 to 2004 we perform the nonparametric Kolmogorov-Smirnov test, the results of which are given in Table 1.7. We conclude that changes over time in $G_{age|k_\tau l_\eta}$ are not significant at none of the above quantile positions for ten countries. Moreover, for Bosnia and Herzegovina, France, Germany, Norway, Portugal, and Slovakia changes in $G_{age|k_\tau l_\eta}$ are significant at only one quantile position. Finally, only in the Czech Republic, Italy, Romania, and Spain, changes in $G_{age|k_\tau l_\eta}$ are significant for most quantile positions.

1.B Appendix: Derivation of the Aggregate Relation in terms of $\bar{K}$ and $\bar{L}$

Let $x_t^j = (k_t^j, l_t^j)'$ denote the observable firm-specific explanatory variables with the corresponding mean vector $\bar{x}_t$. Further, $\Sigma_t = \begin{pmatrix} (\sigma_t^k)^2 & \sigma_t^{kl} \\ \sigma_t^{kl} & (\sigma_t^l)^2 \end{pmatrix}$ denotes the covariance matrix of x_t^j across J_t. According to Hildenbrand and Kneip (2005) the growth rate g_t of the aggregate response variable is given by

$$g_t = \beta_{t-1}'(\bar{x}_t - \bar{x}_{t-1}) + tr[\Delta_{t-1}(\Sigma_t^{1/2}\Sigma_{t-1}^{-1/2} - \mathbb{I})] + \text{ other effects,} \tag{1.17}$$

where $\mathbb{I}$ is the identity matrix, $\beta_{t-1} = (\beta_{t-1}^k, \beta_{t-1}^l)'$ is a vector and $\Delta_{t-1} = \begin{pmatrix} \delta_{t-1}^k & \delta_{t-1}^{kl} \\ \delta_{t-1}^{kl} & \delta_{t-1}^l \end{pmatrix}$ is a matrix of coefficients. Under coordinate-wise standardization (in Assumption 1) Σ_t is replaced by $\tilde{\Sigma}_t = \begin{pmatrix} (\sigma_t^k)^2 & 0 \\ 0 & (\sigma_t^l)^2 \end{pmatrix}$ and the first two rhs terms in (1.17) simplify to

$$\beta_{t-1}^k(\bar{k}_t - \bar{k}_{t-1}) + \beta_{t-1}^l(\bar{l}_t - \bar{l}_{t-1}) + \delta_{t-1}^k(\frac{\sigma_t^k - \sigma_{t-1}^k}{\sigma_{t-1}^k}) + \delta_{t-1}^l(\frac{\sigma_t^l - \sigma_{t-1}^l}{\sigma_{t-1}^l}), \tag{1.18}$$

[18]In fact, when analyzing the evolution of $G_{t,age|k_\tau l_\eta}$ we focus on the distribution of firm age for firms with $k_t^j \in [k_{t,\tau-0.025}, k_{t,\tau+0.025}]$ and $l_t^j \in [l_{t,\eta-0.025}, l_{t,\eta+0.025}]$.

country	$\tau=\eta=0.1$	$\tau=\eta=0.25$	$\tau=\eta=0.5$	$\tau=\eta=0.75$	$\tau=\eta=0.9$
Austria	0.243	0.107	0.616	0.561	0.862
Belgium	0.794	0.703	0.772	0.416	0.978
Bosnia & H.	0.059	0.058	0.827	0.473	0.003*
Bulgaria	0.548	0.980	0.730	0.884	0.382
Czech R.	0.011*	0.956	0.244	0.028*	0.001*
Denmark	0.952	0.999	0.114	0.808	0.723
Estonia	0.149	0.087	0.595	0.781	0.439
Finland	0.597	0.487	0.124	0.422	0.600
France	0.708	0.825	0.740	0.029*	0.996
Germany	0.532	0.032*	0.497	0.977	0.853
Great Britain	0.266	0.546	0.215	0.753	0.235
Italy	0.000*	0.001*	0.000*	0.672	0.474
Netherlands	0.876	0.913	0.720	0.879	0.888
Norway	0.116	0.436	0.373	0.064	0.000*
Poland	0.213	0.083	0.334	0.496	0.998
Portugal	0.499	0.029*	0.121	0.768	0.995
Romania	0.000*	0.001*	0.001*	0.000*	0.000*
Slovakia	0.957	0.021*	0.649	0.974	0.305
Spain	0.000*	0.000*	0.020*	0.000*	0.053
Sweden	0.167	0.679	0.115	0.238	0.464

Asterisks correspond to *p*-values smaller than 0.05 and indicate that changes in the distribution were statistically significant at the 5% level.

Table 1.7: Kolmogorov-Smirnov test of equality of $G_{2003,age|k_\tau l_\eta}$ and $G_{2004,age|k_\tau l_\eta}$ for different quantile positions τ and η.

where

$$\delta^k_{t-1} = \frac{1}{\bar{Y}_{t-1}} \int (k - \bar{k}_{t-1}) \partial_k \bar{f}_{t-1}(k, l, A)\, dG_{t-1,klA}$$

and

$$\delta^l_{t-1} = \frac{1}{\bar{Y}_{t-1}} \int (l - \bar{l}_{t-1}) \partial_l \bar{f}_{t-1}(k, l, A)\, dG_{t-1,klA}.$$

For the sake of comparability with conventional growth models, we are interested in a relationship like (1.17) but in terms of changes in aggregate levels $\bar{K}$ and $\bar{L}$ rather than in terms of aggregate *log* levels $\bar{k}$ and $\bar{l}$. More specifically, we want to arrive at a relationship for the growth rate containing

$$\beta^k_{t-1}(\log \bar{K}_t - \log \bar{K}_{t-1}) + \beta^l_{t-1}(\log \bar{L}_t - \log \bar{L}_{t-1}).$$

We start[19] with the definition of $\log \bar{K}_t$.

$$\log \bar{K}_t = \log \Big[\int K dG_{t,K} \Big] = \log \Big[\int \exp(k) dG_{t,k} \Big]. \tag{1.19}$$

For two periods t and $t-1$ Assumption 1 (Structural stability of G_{kl}) implies

$$G_{t-1,k}\Big(\frac{\sigma_t^k}{\sigma_{t-1}^k}(k - \bar{k}_{t-1}) + \bar{k}_t\Big) = G_{t,k}(k).$$

Hence, we can rewrite (1.19) by

$$\begin{aligned} \log \bar{K}_t &= \log \Big[\int \exp\Big(\frac{\sigma_t^k}{\sigma_{t-1}^k}(k - \bar{k}_{t-1}) + \bar{k}_t\Big) dG_{t-1,k} \Big] \\ &= \bar{k}_t + \log \Big[\int \exp\Big(\frac{\sigma_t^k}{\sigma_{t-1}^k}(k - \bar{k}_{t-1})\Big) dG_{t-1,l} \Big] \end{aligned}$$

Now, we define a function q from $\mathbb{R}_+$ to $\mathbb{R}$ such that

$$q(\sigma^k) := \log \Big[\int \exp\Big(\frac{\sigma^k}{\sigma_{t-1}^k}(k - \bar{k}_{t-1})\Big) dG_{t-1,k} \Big].$$

By the definition of q we have $q(\sigma_t^k) = \log \bar{K}_t - \bar{k}_t$ and simple algebra yields $q(\sigma_{t-1}^k) = \log \bar{K}_{t-1} - \bar{k}_{t-1}$. From these properties of q it follows that

$$\bar{k}_t - \bar{k}_{t-1} = \log \bar{K}_t - \log \bar{K}_{t-1} - [q(\sigma_t^k) - q(\sigma_{t-1}^k)].$$

Further, by the first order Taylor approximation of $q(\sigma^k)$ at σ_{t-1}^k we obtain

$$\begin{aligned} q(\sigma_t^k) &\approx q(\sigma_{t-1}^k) + \partial_{\sigma^k} q(\sigma^k)\big|_{\sigma^k = \sigma_{t-1}^k} \cdot (\sigma_t^k - \sigma_{t-1}^k) \\ &= q(\sigma_{t-1}^k) + \frac{1}{\sigma_{t-1}^k \bar{K}_{t-1}} \int (k - \bar{k}_{t-1}) \exp(k) dG_{t-1,k} \cdot (\sigma_t^k - \sigma_{t-1}^k). \end{aligned}$$

Consequently,

$$\beta_{t-1}^k(\bar{k}_t - \bar{k}_{t-1}) = \beta_{t-1}^k(\log \bar{K}_t - \log \bar{K}_{t-1}) - \frac{\beta_{t-1}^k}{\bar{K}_{t-1}} \int (k - \bar{k}_{t-1}) \exp(k) dG_{t-1,k} \cdot \Big(\frac{\sigma_t^k - \sigma_{t-1}^k}{\sigma_{t-1}^k}\Big).$$

[19] The derivation for $\log \bar{L}_t$ can be carried out analogously.

Doing analogous derivations for $\log \bar{L}_t$, we obtain

$$
\begin{aligned}
g_t &= \beta^k_{t-1}(\log \bar{K}_t - \log \bar{K}_{t-1}) + \beta^l_{t-1}(\log \bar{L}_t - \log \bar{L}_{t-1}) \\
&+ \gamma^k_{t-1}\Big(\frac{\sigma^k_t - \sigma^k_{t-1}}{\sigma^k_{t-1}}\Big) + \gamma^l_{t-1}\Big(\frac{\sigma^l_t - \sigma^l_{t-1}}{\sigma^l_{t-1}}\Big) + \text{ other effects,}
\end{aligned}
$$

where

$$
\gamma^k_{t-1} = \delta^k_{t-1} - \frac{\beta^k_{t-1}}{\bar{K}_{t-1}} \int (k - \bar{k}_{t-1}) \exp(k) dG_{t-1,k}
$$

and

$$
\gamma^l_{t-1} = \delta^l_{t-1} - \frac{\beta^l_{t-1}}{\bar{L}_{t-1}} \int (l - \bar{l}_{t-1}) \exp(l) dG_{t-1,l}.
$$

Chapter 2

Structural Stability of the Joint Distribution of Income and Wealth*

2.1 Introduction

The notion of *structural stability* can be found in many fields of economic research. However, its definition turns out to be different for different fields of research. From the econometric point of view, for example, one could regard a postulated model to be structurally stable, if no structural breaks occur in the sense that parameter values are assumed to be constant over time, see e.g. Chow (1983) or Hansen (1992). A slightly different definition is used in game theory, where a game is considered to satisfy the property of structural stability, if small perturbations of the payoff matrix do not alter the qualitative nature of the outcome, see e.g. Palis and Smale (1970). In this chapter we will confine ourselves to the notion of structural stability in the context of aggregation theory.

The concept of structural stability has been present in aggregation theory since the papers of Malinvaud (1993).[1] Unlike typical macroeconomic models that link aggregate response to aggregate explanatory variables, Malinvaud's idea was to model aggregates in terms of the entire joint distribution of all individual variables. This distribution was assumed to have a certain parametric form (structure), e.g., the log-normal distribution in case of the distribution of income or the firm size. In

*This chapter is based on Paluch (2004).

[1] Malinvaud (1993) was in the main the English translation of his paper in French from 1956.

modeling changes over time in this distribution, he made use of the empirical fact that its structure does not change over time, i.e., the log-normality prevails, and its entire evolution can be well captured by changes in only few of its parameters like the mean or the variance. It is this phenomenon which Malinvaud refers to as structural stability.[2]

In fact, the concept of structural stability as stated by Malinvaud (1993) cannot be applied to distributions which are poorly approximated by a parametric form.[3] If one does not want to impose any assumptions on the parametric form of the analyzed distributions, one is forced to find a more flexible (nonparametric) counterpart of Malinvaud's idea. Instead of keeping the parametric structure constant and allowing for changes in few parameters, one can fix the values of some parameters and allow the shape of the distribution to vary. This can be achieved by simple transformations of the distribution like centering, scaling or standardizing. This concept has been formulated by Hildenbrand and Kneip (1999). Their definition of structural stability of a sequence of distributions states that, by applying a simple transformation to the original distribution, the local time-invariance of the sequence of transformed distributions can be achieved. Hence, the local time-invariance holds if the period-to-period changes in the sequence of transformed distributions can be regarded as statistically insignificant. Therefore, if a transformed distribution turns out to be locally time-invariant, the complicated evolution of the original distribution can be captured completely by the changes in the parameters used for the transformation.[4]

The most important implication of structural stability is the possibility to predict the shape of the future distributions. Indeed, if structural stability holds, the original distribution in period $t + 1$ is completely determined by the original distribution in

[2]This empirical regularity has been mentioned not later than in the 19th century for the case of income distributions by Pareto (1896-1897).

[3]The assumption of the log-normality of the income distribution is violated for variety of countries because of its multimodality.

[4]Consequently, one can distinguish several versions of structural stability depending on the strictness of this assumption, e.g. the local time-invariance of a standardized distribution is a weaker assumption than the corresponding assumption for the centered or relative distribution.

period t and the parameters, like the mean or the variance, which have been used for transformation, in period $t+1$. As a consequence, the very complex modeling of the short-run evolution of this distribution can be reduced to the modeling of changes in the parameters. Interestingly, despite the arising new possibilities of modeling aggregate behavior on the basis of structural stability, one can hardly find applications of this concept in the literature.[5] Indeed, to the author's knowledge, there is only one theory that models aggregation under structural stability. In order to model a relative change in an aggregate in an economy, Hildenbrand and Kneip (1999 and 2005) propose an approach based on the evolution over time of distributions of observed and unobserved explanatory variables.

Surprisingly, even in the empirical literature the explicit verification of structural stability is very seldom. For example, the evolution of individual or cross-country relative income distribution has been studied extensively in the economic literature. Empirical work on this topic, e.g. Cowell, Jenkins and Litchfield (1996), Quah (1997) or Sala-i-Martin (2002), however, was targeted mainly at the aspect of changing inequality, poverty, and convergence of these distributions.[6] Indeed, we are aware of only two papers that studied empirical validity of structural stability of the distribution of households' income. In Hildenbrand, Kneip and Utikal (1999), graphical analysis of the evolution of relative and standardized income distribution for Great Britain is presented. It turns out that simple transformations of this distribution like scaling or standardizing can remove a huge part of its variation over the years. Pittau and Zelli (2001) analyze trends in income distribution in Italy both graphically and

---

[5]Schumpeter (1951), as cited by Malinvaud (1993), regrets that researchers do not exploit the potentialities of structural stability:

"Few if any economists seem to have realized the possibilities that such invariants hold out for the future of our science... nobody seems to have realized that the hunt for, and the interpretation of, invariants of this type might lay the foundations of an entirely novel type of theory"

[6]The mentioned papers apply kernel density estimation and are therefore not the typical ones in the empirical literature on convergence and changing inequality of the income distribution. Usually, the analysis of these issues is based solely on the study of the changes in the characteristic parameters of this distribution, like the Gini-coefficient, variance of log-income, Atkinson (1970) indices or the mean-median ratio. One example of papers following this approach is Gottschalk and Smeeding (2001) that contains an international comparison of the income inequality and its changes over time.

by means of a statistical test and show that the distribution of relative incomes is locally time-invariant for many periods.

The aforementioned empirical studies concerned only univariate distributions. However, in the formulation of structural stability, Malinvaud mentions the *joint distribution* of all individual exogeneous variables. This motivates our work, which extends the empirical study of Hildenbrand et al (1999) on income distribution in two aspects. First, we incorporate an additional variable, namely wealth of a household. Consequently, in this chapter we will study the short-run dynamics of the joint distribution of households' income and wealth. In particular, we try to find local time-invariance in this distribution after exposing it to scaling or standardizing transformations. Second, to endorse graphical arguments and to check whether the observed changes over time in this distribution are statistically significant, a nonparametric time-invariance test as suggested by Li (1996) is performed.

The remainder of this chapter is organized as follows. We give a motivation for the study of the joint distribution of income and wealth and its evolution in Section 2.2. A brief description of one particular application of the aggregation model formulated by Hildenbrand and Kneip (2005) with emphasis on the hypothesis of structural stability is given. In Section 2.3 we present the data from the Family Resources Survey used in our empirical analysis and report some descriptive statistics of the underlying population of British households. Furthermore, we describe the econometric methods which are employed in this work to analyze the short-run dynamics of distributions. Finally, we look for a transformation of the original distribution that is sufficient to yield the local time-invariance of the resulting distribution in Section 2.4. A short summary and conclusions are provided in Section 2.5.

2.2 A Motivating Example: Aggregation of Households' Consumption Expenditure

The aim of the aggregation model in Hildenbrand and Kneip (2005) is to explain the relative change in an aggregate over time. The starting point of this model is the behavioral relation of the microunit, which links explanatory variables to the individual response variable. The modeling occurs, amongst others, in terms of changes in the distribution of *observable* and *unobservable* individual exogeneous variables across the whole population. In particular, the joint distribution of all *observable* micro-specific variables across the whole population is assumed to be structurally stable.

As already mentioned in the Introduction, one application of the model stated in Hildenbrand and Kneip (2005) is the aggregation of households' consumption expenditures. For this particular case, the whole population in period t - denoted by H_t - consists of households h, who have to decide about the level of their consumption expenditure. Therefore, their behavioral relation links following explanatory variables: income, wealth, prices, interest rates, preference parameters of the utility function, expectational variables like expected future income, life expectancy etc. to the response variable, i.e., the consumption expenditure of a household. The consumption theoretical application presented in Hildenbrand and Kneip (2005) treats only two of the variables mentioned above as observable[7] and micro-specific. These two variables are the household's income and wealth denoted by y_1^h and y_2^h, respectively, and are captured in the vector of observable micro-specific variables of household h, which is denoted by y^h. Consequently, for this particular application of the model, the joint distribution of income and wealth across the whole population, denoted by $\text{distr}(y \mid H_t)$, is assumed to evolve in the structurally stable way. Hildenbrand and Kneip (2005) state this assumption in terms of the the standardized distribution, i.e.,

Hypothesis: Structural stability of $\text{distr}(y \mid H_t)$

The standardized joint distribution of log-income and log-wealth across the whole population changes sufficiently slowly over time in the sense that this distribution can be considered as approximately equal for two periods that are close to each other.

---

[7]The main criterion to consider a variable to be observable is the availability of the data on this variable. It is often the case that even if the variable is observable in reality, e.g. some aspects of wealth, households are either not asked for or they just do not know its exact value.

In the empirical part of this chapter we will study the evolution of the relative and standardized joint distribution of log-income and log-wealth. Therefore, the empirical results can be used to verify the hypothesis of structural stability of the joint distribution of log-income and log-wealth as formulated above by Hildenbrand and Kneip (2005).

2.3 Data Treatment and Methodology

Our empirical analysis is based on cross-sectional data from the British Family Resources Survey (henceforth referred to as FRS). This survey was started in 1992 by order of the Department of Social Security. For each individual in the household it collects information on income, savings and financial assets and on a variety of socio-economic and demographic variables like age or employment status of each household's member. Each year about 25,000 households are interviewed. The information gained by this survey is mainly used by non-governmental organizations to simulate and analyze the response of the population to new policy measures. Furthermore, basically due to the large sample sizes, the FRS data is gaining popularity in empirical research being a reliable basis for studies on dynamics of income and wealth, see e.g. Piachaud and Sutherland (2002) or Ginn and Arber (2000).

The variables used for the search of structural stability are income and financial wealth. Unfortunately, due to inconsistency problems in the definitions of these two variables, the time horizon for the analysis had to be reduced to six years, i.e. 1996-2001. As we look for local and not global time-invariance of the distribution, the span of only six years data is adequate for analysis.

The income variable used in this work is household's weekly disposable non-property income, which is defined as the intra household sum of total net earnings from all sources (excluding property income), net pensions and various state transfers like benefit income, income in kind, etc. As far as financial wealth is concerned, balances from following accounts are included: current accounts, savings accounts, gilts, trusts, stocks, shares, national saving certificates, save-as-you-earn contributions, yearly plans, premium bonds, pensioner guaranteed income bonds, etc., whereas life insurance is not included. The value of household's financial wealth is obtained in the following way. At the beginning of the interview about household's wealth, the head of family is asked whether its total amount of capital is between £1500 and £20000. Should it lie within this interval, further questions regarding the composition and amount of financial wealth are asked. Otherwise, the amount of capital is approximated by dividing the yearly investment income from aforementioned accounts by the corresponding account specific interest rates.

It is a well known empirical fact that the distributions of income and wealth are right-skewed. The analysis of the time-invariance of a distribution is much simpler if it is symmetric, because such a distribution can be more easily characterized by its moments like mean, variance, etc. Furthermore, at the outset of our empirical study, the large changes in the distributions of income and wealth can be noticeably reduced by using logarithmic transformation. Therefore, for the analysis in this chapter we use the log-values of income and financial wealth. The desired effect achieved by the logarithmic transformation can be seen in Figure 2.1, where the kernel density estimates of the distributions of income and log-income for years 1996-2001 are plotted.

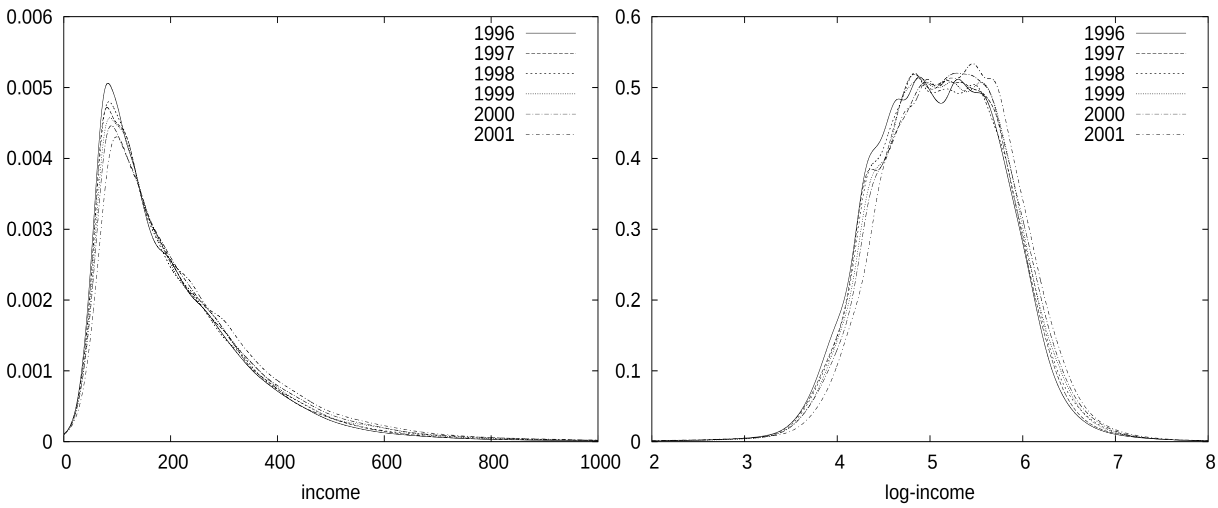

Figure 2.1: Kernel density estimates of income and log-income distributions across H_t for 1996-2001.

However, the verification of the hypothesis of structural stability of the joint distribution of log-income and log-wealth creates the following problem. Typically, not all households hold financial assets. Because of the use of log-values of income and wealth, the joint distribution distr$(y \mid H_t)$ is defined only for strict positive values of y. This forces us to conduct a separate analysis for subpopulation H_t^1 containing all households in the population H_t with positive wealth[8] and subpopulation H_t^0, which contains the remaining households in the population. Interestingly, the relative size of H_t^1, i.e. H_t^1/H_t, does not change substantially over time. The descriptive statistics for the whole population H_t and the coefficient of correlation between log-income and log-wealth across H_t^1 are given in Table 2.1.

year	group size			mean log-income		mean	corr.
	H_t^0	H_t^1	H_t^1/H_t	H_t^0	H_t^1	log-wealth	
1996	9401	16019	63.01%	4.832 (0.587)	5.230 (0.716)	7.979 (1.671)	0.105
1997	8911	14387	61.75%	4.870 (0.596)	5.255 (0.725)	7.848 (1.658)	0.075
1998	8816	13951	60.65%	4.884 (0.591)	5.270 (0.733)	7.848 (1.649)	0.097
1999	9895	14929	60.13%	4.929 (0.589)	5.288 (0.737)	7.899 (1.689)	0.079
2000	9763	13813	58.58%	5.061 (0.674)	5.243 (0.720)	7.914 (1.677)	0.065
2001	10196	14931	59.42%	5.014 (0.630)	5.367 (0.716)	7.805 (1.606)	0.067

Terms in parentheses are standard deviations of log-values.

Table 2.1: Descriptive statistics and the coefficient of correlation between log-income and log-wealth across H_t^1.

As far as econometric methods applied in this work are concerned, all distributions have been estimated nonparametrically using the adaptive bandwidth kernel density estimator with the second order Gaussian kernel function. The pilot bandwidth was chosen according to Sheather and Jones (1991) plug-in method.

Once densities are estimated, an important question arises, whether the observed changes over time in the estimates are statistically significant. In order to answer this question, we apply a nonparametric test of closeness between two distribution

[8]We treat all household with the capital amount of less than £100 (in prices of 1988) as if they had no wealth. This is motivated by the fact that for each household that claims its financial wealth to be less than £1500, the value of financial wealth is approximated by the division of household's yearly investment income by the interest rate. The breaking point of £100 corresponds to the negligible household's weekly investment income of £0.10 if one assumes that the interest rate is at 5%.

functions as proposed by Li (1996). Given the observations[9] $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ drawn from the corresponding unknown density functions f_X and f_Y the test is based on the integrated squared difference between f_X and f_Y denoted by I and defined by

$$\begin{aligned} I &= \int [f_X(t) - f_Y(t)]^2 dt = \int [f_X^2(t) + f_Y^2(t) - 2f_X(t)f_Y(t)]dt \\ &= \int f_X(t)dF_X(t) + \int f_Y(t)dF_Y(t) - 2\int f_Y(t)dF_X(t). \end{aligned}$$

In our work the densities f_X and f_Y correspond to the distributions from different time periods, e.g. f_X and f_Y are the relative log-income distributions in period t and $t+1$ respectively. The feasible estimator of I, denoted by I_n, can be obtained, if one substitutes the density functions f_X and f_Y by their kernel estimates $\hat{f}_X$ and $\hat{f}_Y$, i.e.,

$$\hat{f}_X(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \quad \text{and} \quad \hat{f}_Y(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - Y_i}{h}\right).$$

Using these estimates and replacing F_X and F_Y by their empirical distribution functions, one can write $I_n = I_{1n} + I_{2n}$, where

$$I_{1n} = \frac{2K(0)}{nh} - \frac{2}{n^2 h}\sum_{i=1}^{n} K\left(\frac{X_i - Y_i}{h}\right) = c(n) + \mathcal{O}(n^{-1})$$

and

$$I_{2n} = \frac{1}{n^2 h}\sum_{i=1}^{n}\sum_{\substack{i \neq j \\ j=1}}^{n}\left[K\left(\frac{X_i - X_j}{h}\right) + K\left(\frac{Y_i - Y_j}{h}\right) - K\left(\frac{Y_i - X_j}{h}\right) - K\left(\frac{X_i - Y_j}{h}\right)\right].$$

The test structure is as follows:

H_0: $f_X(x) = f_Y(x)$ almost everywhere

H_1: $f_X(x) \neq f_Y(x)$ for some x.

Under the null hypothesis of time-invariance and assuming that for $h \to 0$ and

[9] For the sake of simplicity of the presentation, we assume the samples of observations on X and Y to be of equal sizes and to be drawn from univariate densities f_X and f_Y. However, the extension of the test for the case of different sample sizes and multivariate distributions is easy. Furthermore, the random variables X and Y need not to be independent in the sense that the possible dependence does not change the asymptotic distribution of the test statistic.

$nh \to \infty$, Li (1996) has shown that $T_n := nh^{1/2}\frac{I_n - c(n)}{\hat{\sigma}_0} \to^d N(0,1)$, where

$$\hat{\sigma}_0 = \frac{2}{n^2 h}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[K\left(\frac{X_i - X_j}{h}\right) + K\left(\frac{Y_i - Y_j}{h}\right) + 2K\left(\frac{X_i - Y_j}{h}\right)\right]\left[\int K^2(u)du\right]$$

and $c(n) = 2K(0)/nh$.

The asymptotic distribution of the test statistic T under the null hypothesis has a slow rate of convergence to the the standard normal distribution. In order to account for this finite sample bias, we perform the bootstrap procedure to approximate the distribution of T. We repeat a following procedure 500 times: Out of the pooled sample $\{X_1, \ldots, X_{n_1}; Y_1, \ldots, Y_{n_2}\} =: \{Z_1, \ldots, Z_{n_1+n_2}\}$ two samples, $\{X_1^*, \ldots, X_{n_1}^*\}$ and $\{Y_1^*, \ldots, Y_{n_2}^*\}$, are randomly drawn with replacement. Then, based on the new samples the test statistic $T_{n,i}^*$ is computed. The empirical distribution of T under the null hypothesis is then estimated from the sample $\{T_{n,1}^*, \ldots, T_{n,500}^*\}$. The bandwidth for testing purposes was obtained as an optimal bandwidth for density estimation for the pooled sample $\{Z_1, \ldots, Z_{n_1+n_2}\}$ according to the Sheather and Jones (1991) plug-in method. A proof of consistency of this bootstrap in the context of testing our hypotheses can be found in Li, Maasoumi and Racine (2007).

2.4 Empirical Results

2.4.1 The Evolution of the Relative Joint Distribution of Log-income and Log-wealth

The relative joint distribution of log-income and log-wealth across the population H^1 in period t is defined as the distribution of $\hat{y}_t^h = (\hat{y}_{t,1}^h, \hat{y}_{t,2}^h) := (y_{t,1}^h/m_{t,1},\, y_{t,2}^h/m_{t,2})$, where $m_{t,1}$ and $m_{t,2}$ denote the mean log-income and mean log-wealth across H_t^1, respectively. For the population H^0 the relative joint distribution of log-income and log-wealth is just the univariate distribution of relative log-income. Mean-scaling of the distribution implies the first moment of the resulting relative distribution to be constant over time and equal to 1. Therefore, one can regard the relative distribution as a detrended one in which only higher moments like variance, skewness or kurtosis may change over time.[10] Consequently, if the shape of the relative distribution does not change significantly over time, the evolution of the original distribution is captured entirely by the changes over time in its mean.

Population H^1

Figures 2.2 and 2.3 show the kernel density estimates of $\mathrm{distr}(\hat{y} \mid H^1_{1996})$ and the associated density contours for years 1996 and 1997, respectively. As one can see in Figure 2.3, the density contours for these two years do not differ noticeably from each other. We have observed this feature also for other years of the sample. This fact can be seen more clearly on two dimensional graphs of marginal distributions of $\mathrm{distr}(\hat{y} \mid H^1_t)$, i.e., the relative log-income distribution and relative log-wealth distribution across H^1_t, which are presented in Figure 2.4.

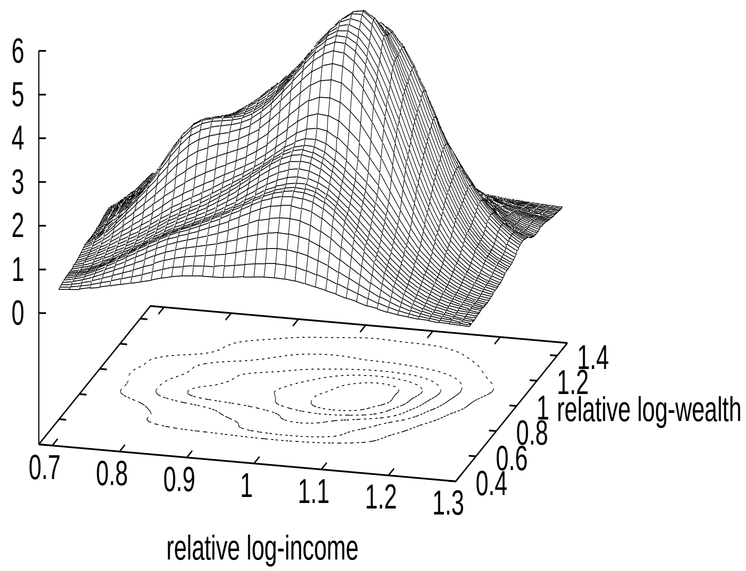

Figure 2.2: Kernel density estimate of $\mathrm{distr}(\hat{y} \mid H^1_{1996})$.

[10]Pittau and Zelli (2001) use a different definition of the relative distribution, which is derived by dividing all observations by the sample median and not the mean. Note that in the case of median-scaling, the mean of this kind of relative distribution will not usually be not time-invariant.

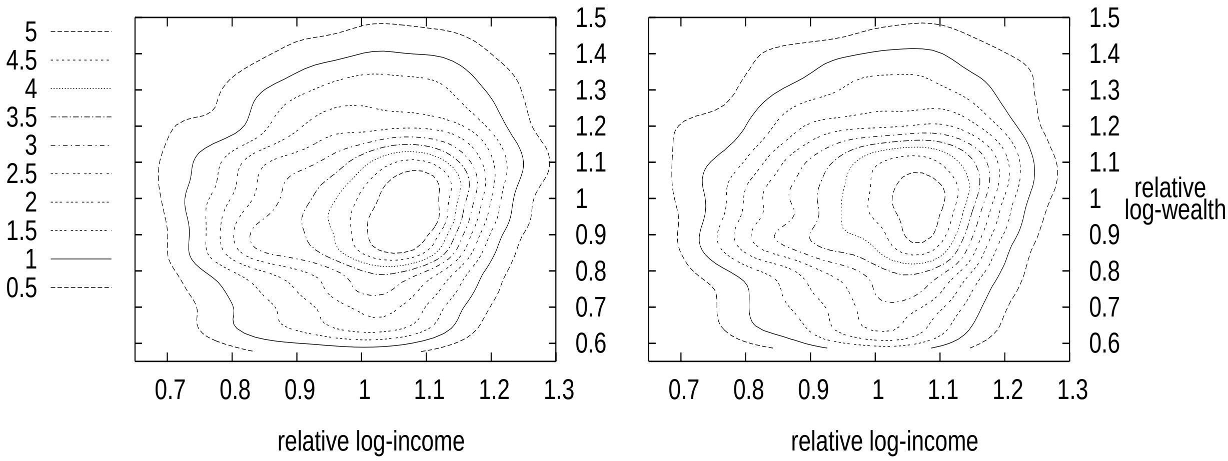

Figure 2.3: Density contours of $\mathrm{distr}(\hat{y} \mid H^1_{1996})$ (left) and $\mathrm{distr}(\hat{y} \mid H^1_{1997})$ (right).

Population H^0

The relative log-income distribution across H^0, which is plotted in Figure 2.5, can be also regarded as stable over time. However, a huge increase in the dispersion of the original distribution in the year 2000 that can be seen in Table 2.1 is reflected in the estimate, which is quite different from that for other years. As the mean-scaling transformation does not account for changes in the dispersion, we can expect the changes during the transitions 1999-2000 and 2000-2001 to be highly significant.

Li (1996) Test Results for the Relative Distributions

The question, whether the observed year-to-year changes are significant or not, cannot be answered without applying proper statistical test. Therefore, in order to study the significance of changes in the relative joint distribution of log-income and log-wealth over time, we apply the Li (1996) test. The test results are given in Table 2.2.

As one can see in Table 2.2, the null hypothesis of equality of $\mathrm{distr}(\hat{y} \mid H^1_t)$ and $\mathrm{distr}(\hat{y} \mid H^1_{t+1})$ cannot be rejected for only one transition period, 1997-1998, which implies that the evolution of $\mathrm{distr}(y \mid H^1)$ is too complex to be captured by only its first moment. As far as the distribution $\mathrm{distr}(\hat{y}_1 \mid H^0)$ is concerned, one cannot reject the equality hypothesis for only two transition periods, 1997-1998 and 1998-1999.

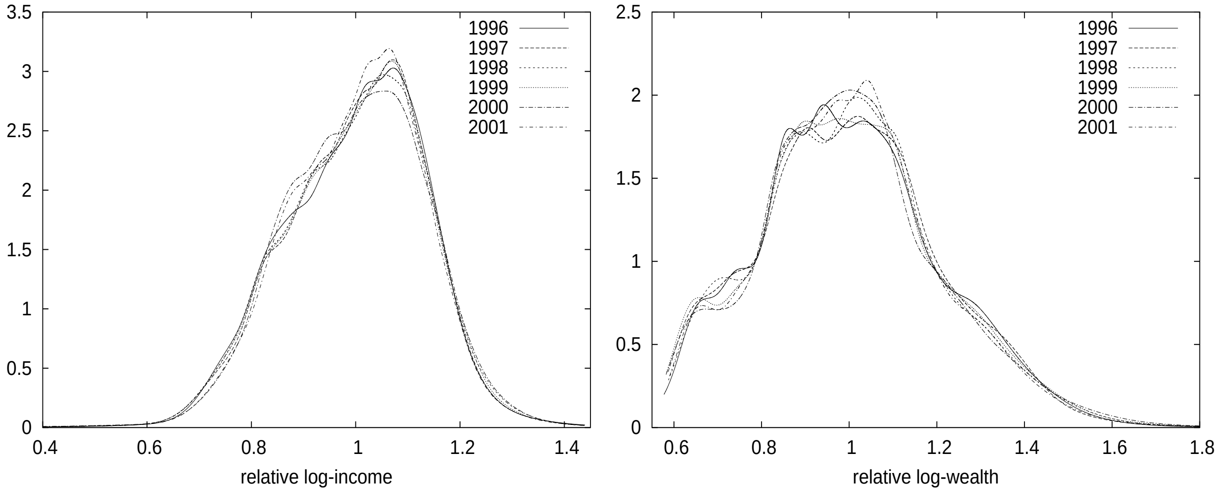

Figure 2.4: Kernel density estimator of the relative log-income distribution and the relative log-wealth distribution across H^1 for 1996-2001.

This motivates the attempt to incorporate further parameters that would account for changes in the dispersion of the original distribution. The most intuitive candidates for this are the elements of the covariance matrix of the original distribution. In the next subsection, we will study the case of standardizing transformation as an example of such an extension.

	Subpopulation H^1		Subpopulation H^0	
transition period	T-stat	empirical p-value	T-stat	empirical p-value
1996 vs. 1997	3.934	0*	2.868	0.004*
1997 vs. 1998	1.061	0.107	-1.054	0.807
1998 vs. 1999	3.173	0*	-1.299	0.902
1999 vs. 2000	12.880	0*	17.354	0*
2000 vs. 2001	6.069	0*	14.816	0*

Asterisk indicate that equality is rejected at the 5% level.

Table 2.2: Li (1996) test results for the distributions distr($\hat{y} \mid H^1$) and distr($\hat{y} \mid H^0$) for 1996-2001.

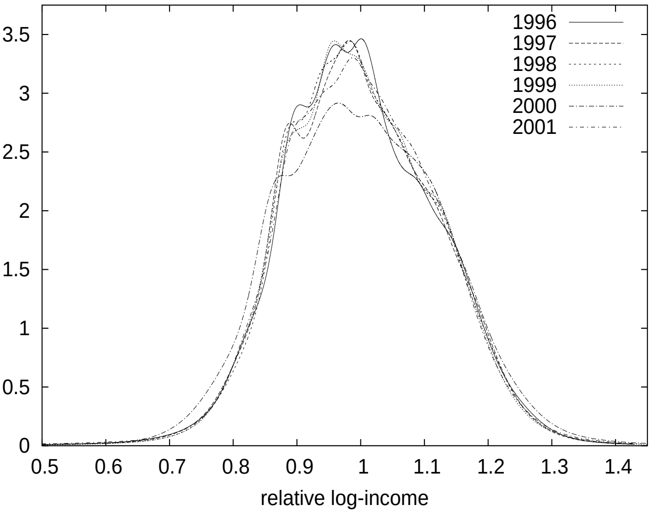

Figure 2.5: Kernel density estimate of the relative log-income distribution across H^0 for 1996-2001.

2.4.2 The Evolution of the Standardized Joint Distribution of Log-income and Log-wealth

The standardized joint distribution of log-income and log-wealth across H^1 in period t is defined as the distribution of $\tilde{y}_t^h := \Sigma_t^{-1/2}(y_t^h - m_t)$, where m_t denotes the vector of means of log-income and log-wealth and Σ_t is the covariance matrix of log-income and log-wealth across H_t^1. The correlation between log-income and log-wealth across the population H^1 presented in Table 2.1 is very small. Therefore, one can approximate this distribution by applying to the original distribution – $\text{distr}(y \mid H_t^1)$ – the simpler version of the standardization, so called coordinate-wise standardization. The coordinate-wise standardized distribution of y_t^h is then defined as the distribution of $(\bar{\tilde{y}}_{t,1}, \bar{\tilde{y}}_{t,2}) := \left(\frac{y_{t,1}-m_{t,1}}{\sigma_{t,1}}, \frac{y_{t,2}-m_{t,2}}{\sigma_{t,2}}\right)$, where $\sigma_{t,1}$ and $\sigma_{t,2}$ denote the standard deviations of log-income and log-wealth, respectively and m_t is the vector of corresponding means across the population H_t^1.

We expect changes over time in the shape of the standardized distribution to be less significant as the corresponding changes in the relative distribution. This is due to the fact that the standardizing transformation (even the coordinate-wise one) implies not only the time-invariance of the vector of means (equal to 0) of the transformed distribution, but also the time-invariance of the variances (equal to 1) of its marginal distributions.

Population H^1

Kernel density estimates of distr($\bar{\tilde{y}} \mid H^1_{1996}$) and the associated density contours for years 1996 and 1997 are presented in Figures 2.6 and 2.7, respectively. As in the case of the relative distribution, the density contours for these years do not change much over time, which also holds for other years. Marginal distributions of distr($\bar{\tilde{y}} \mid H^1_t$), i.e. the standardized log-income distribution and the standardized log-wealth distribution across H^1_t are presented in Figure 2.8 and reveal small variations in these distributions.

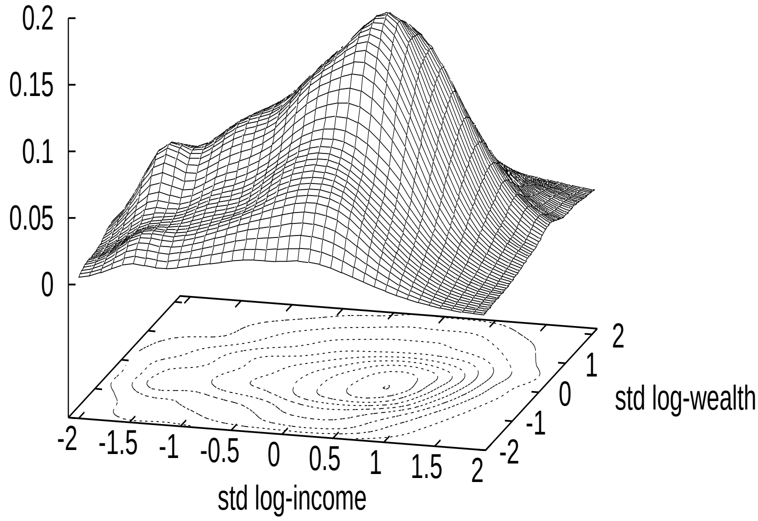

Figure 2.6: Kernel density estimate of distr($\bar{\tilde{y}} \mid H^1_{1996}$).

Population H^0

Figure 2.9 comprises the evidence for the strength of structural stability in showing how even considerably different original distributions can be transformed to very similar ones by controlling for changes in only few parameters. The original distribution of log-income for the year 2000 differs much from that for other years, however, if one applies standardization, the resulting distributions are very similar for all years.

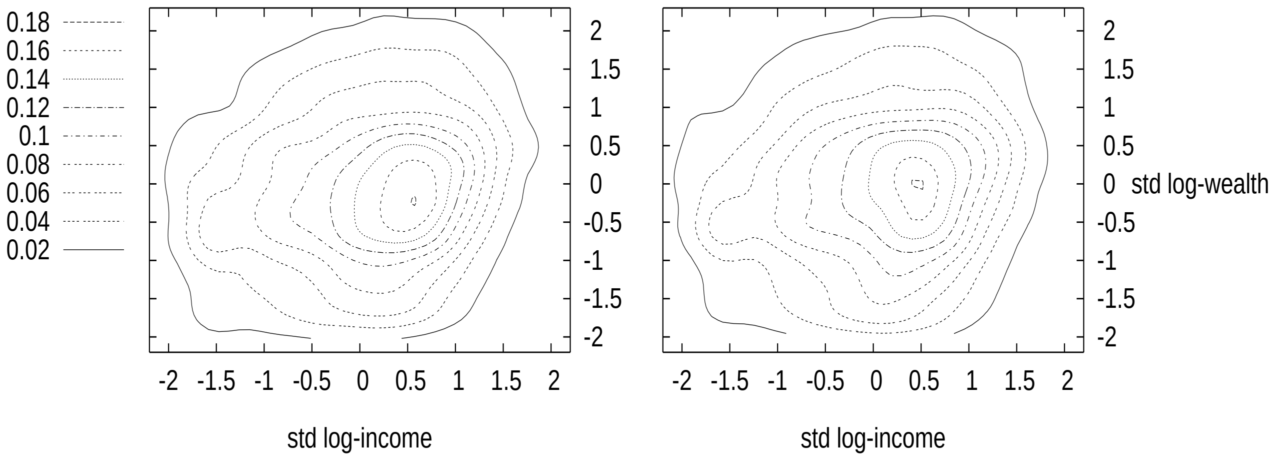

Figure 2.7: Density contours of $\mathrm{distr}(\bar{\tilde{y}}\,|\,H^1_{1996})$ (left) and $\mathrm{distr}(\bar{\tilde{y}}\,|\,H^1_{1997})$ (right).

Note that this is in contrast to the case of the corresponding relative distributions as shown in Figure 2.5.

	Subpopulation H^1		Subpopulation H^0	
transition period	T-stat	empirical p-value	T-stat	empirical p-value
1996 vs. 1997	0.182	0.392	2.354	0.011*
1997 vs. 1998	0.073	0.468	-1.372	0.912
1998 vs. 1999	0.062	0.457	-0.945	0.715
1999 vs. 2000	0.160	0.391	2.004	0.017*
2000 vs. 2001	0.199	0.344	4.107	0**

Asterisks * (**) indicate the rejection of equality at the 5% (1%) level.

Table 2.3: Li (1996) test results for the distributions $\mathrm{distr}(\bar{\tilde{y}}\,|\,H^1_t)$ and $\mathrm{distr}(\tilde{y}\,|\,H^0_t)$ for 1996-2001.

Li (1996) Test Results for the Standardized Distribution

The null hypothesis of equality of $\mathrm{distr}(\bar{\tilde{y}}|\,H^1_t)$ and $\mathrm{distr}(\bar{\tilde{y}}\,|\,H^1_{t+1})$ cannot be rejected for all years within the time period 1996-2001. These results, given in Table 2.3, indicate the possibility of capturing the evolution of the entire distribution $\mathrm{distr}(\bar{\tilde{y}}\,|\,H^1_t)$ by only few parameters, namely the means and the standard deviations. As for the

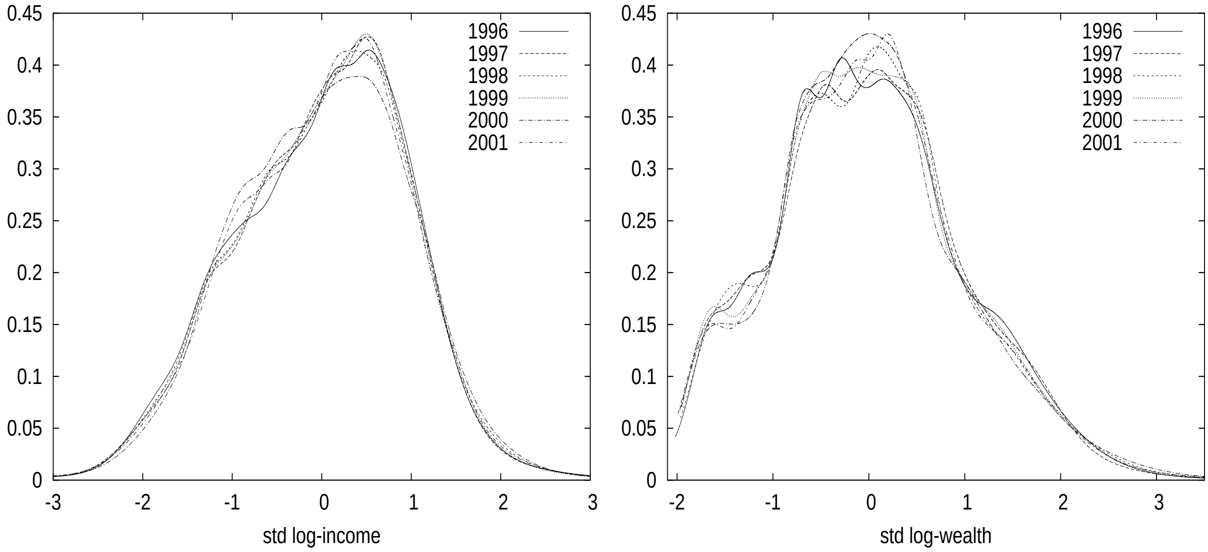

Figure 2.8: Kernel density estimate of the standardized log-income distribution and the standardized log-wealth distribution across H^1 for 1996-2001.

population H^0, the hypothesis of equality cannot be rejected at the 5% significance level for the transitions 1997-1998 and 1998-1999. Further, one cannot reject the equality at the 1% level for the transitions 1996-1997 and 1999-2000. The changes in the standardized distribution of log-income between 2000 and 2001 turn out to be statistically significant at the 1% level.

2.5 Summary and Conclusions

The main aim of this chapter was to examine the short-run dynamics of the joint distribution of income and wealth of British households on the basis of the Family Resources Survey 1996-2001. The focal point of our analysis is the property of structural stability of this distribution – a notion that was formulated firstly by Malinvaud (1993) for distributions of a certain parametric form and was reformulated for the nonparametric case by Hildenbrand and Kneip (1999). In this work, we want to avoid any assumptions on the shape of this distribution and we follow the latter approach. According to this concept, if a sequence of distributions can be exposed to a simple transformation in that manner that the sequence of the transformed distributions is

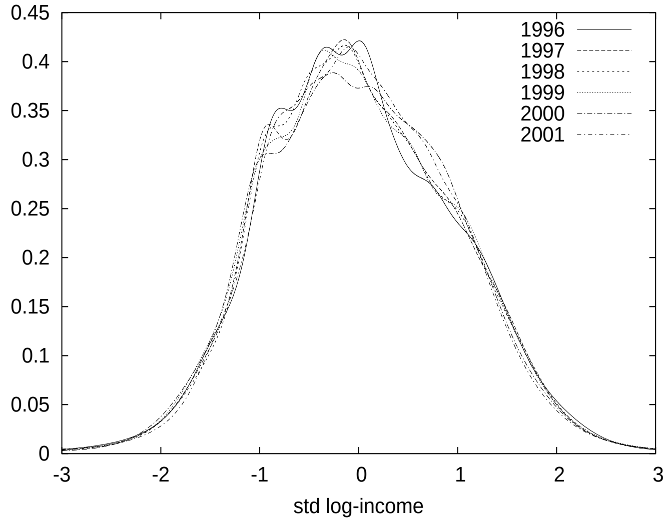

Figure 2.9: Kernel density estimate of the standardized log-income distribution across H^0 for 1996-2001.

locally time-invariant, then the sequence of original distributions is said to be structurally stable. In our search for a simple transformation of a original distribution, i.e. the joint distribution of log-income and log-wealth, that yields local time-invariance of the transformed distribution, we analyze two transformations. The first one, mean-scaling, which could control for the changes over time in mean log-income and mean log-wealth and resulted in the relative joint distribution of log-income and log-wealth, was not sufficient to support the hypothesis of structural stability. However, after applying the standardizing transformation, which accounted for changes in means and dispersion of the original distribution we obtained a sequence of distributions that was local time-invariant, i.e. the period-to-period changes in this sequence were statistically insignificant for almost all years in our sample. This fact empirically supports the hypothesis of structural stability of the joint distribution of income and wealth providing a justification for using this hypothesis in theoretical aggregation models such as the model in Hildenbrand and Kneip (2005).

Chapter 3

Individual versus Aggregate Income Elasticities for Heterogeneous Populations*

3.1 Introduction

Long before the formal economic theory of consumer behavior (and the concept of a demand function) was developed, it was recognized that income is an important explanatory variable for consumer demand. We refer to Stigler (1954) for the early history of empirical studies. Certainly, there are other explanatory variables, such as prices and preferences. In order to derive a complete set of explanatory variables, one needs a precise and complete description of the decision situation. Does the consumer face an atemporal or intertemporal decision with or without uncertainty?

These alternative decision problems are studied in detail in the microeconomic theory of consumer behavior (e.g., Deaton and Muellbauer, 1980, Romer, 2006; for a concise formulation see Section 5.2 of Hildenbrand and Kneip, 2005). For a given period (e.g., a specified year) this leads to a relation which can be generally written in the following form:

$$c^h = f(x^h, v^h),$$

where c^h denotes the expenditure in current prices on a certain category of consumption goods (such as food or services) of consumer h, x^h is disposable income, and v^h denotes the vector of all other explanatory variables. The nature of variables subsumed by v^h crucially depends on the decision situation. In any case v^h will contain

*This chapter is based on Paluch, Kneip and Hildenbrand (2007).

prices and preference parameters. In an inter-temporal setting v^h will also incorporate suitably formalized future expectations.

In order to measure how sensitive consumer h reacts to an income change under the ceteris paribus condition that v^h remains constant, one considers the elasticity of consumption expenditure with respect to income, 'income elasticity' for short, defined by

$$\beta(x^h, v^h) := \frac{x^h}{c^h} \partial_x f(x^h, v^h) = \partial_y \log f(e^{y^h}, v^h),$$

where $y^h = \log x^h$. Thus, if the consumer's income increases by one percent, his consumption expenditure increases by β percent.

For economic policy analysis one needs mean (aggregate) consumption expenditure across a large and heterogeneous population H of households. Let $\nu_{x,v}$ denote the joint distribution of the explanatory variables x^h and v^h across the population H. Then mean consumption expenditure is equal to $C_{mean} := \int f(x,v) d\nu_{x,v} \equiv F(\nu_{x,v})$. Therefore the 'explanatory variable' for mean demand is the distribution $\nu_{x,v}$. The marginal ν_x is the income distribution, and $X_{mean} := \int x d\nu_x$ is mean income.

Consider a change in income on the micro-level, $x^h \to \tilde{x}^h$, under the above ceteris paribus condition. This leads to a new distribution $\tilde{\nu}_{x,v}$ and a changed mean income $\int x d\tilde{\nu}_x$. If, for a heterogeneous population, one wants to relate the resulting change in mean consumption expenditure to the change in mean income, then one has to specify, either on the micro-level how the change in mean income is allocated across the households in the population, or on the distributional level, how the changed distribution $\tilde{\nu}_{x,v}$ is generated from $\nu_{x,v}$.

Throughout this chapter we consider a proportional change on the distributional level (a precise definition is given in Section 3.2). This corresponds to the concept of "mean scaled" income distributions as introduced by Lewbel (1990, 1992). Thus, relative income distributions remain unchanged and, hence, income inequality measures such as Gini or the coefficient of variation are unaffected.

In this setup the elasticity of mean consumption expenditure with respect to mean income, 'aggregate income elasticity' for short, is given by (see Section 3.2)

$$\beta_{agg} = \frac{X_{mean}}{C_{mean}} \partial_\mu \left(\int f(\frac{\mu}{X_{mean}} x, v) d\nu_{x,v} \right) \Big|_{\mu = X_{mean}}.$$

Hence, if, for example, the income of *each* consumer increases by one percent, then mean consumption expenditure increases by β_{agg} percent. We emphasize that β_{agg} will in general not be equal to the mean of individual elasticities. Indeed, a simple calculation given in Section 3.2 leads to

$$\beta_{agg} = \beta_{mean} + \frac{1}{C_{mean}} \mathrm{Cov}(f(x,v), \beta(x,v)),$$

where β_{mean} denotes the mean of individual elasticities, and $\mathrm{Cov}(f(x,v), \beta(x,v))$ is the covariance between individual consumption expenditures and individual elasticities with respect to the distribution $\nu_{x,v}$.

What can be said about the sign or the magnitude of the covariance term? Do households with large demand tend to have large or small elasticities? Under which circumstances can one expect the covariance term to be negligible? The latter is often implicitly assumed in applied work when the magnitude of the estimated aggregate elasticity is interpreted in terms of individual behavior.

Even in the case of a population which is homogeneous in demand behavior, without specifying the demand function, nothing definitive can be said about the sign of the covariance term. For example, even if the common demand function describes the demand for a necessity the sign of the covariance term can be positive or negative. Consequently, the above questions have to be answered by empirical studies.

Many contributions in the literature estimate elasticities based on cross-section and panel data. The standard approach relies on a parametric modeling of demand or consumption expenditure. In addition to specifying the functional relationship between consumption c^h, income x^h and prices p in the current period, possible dependencies between income level x^h and all remaining household specific explanatory variables in v^h have to be taken into account in parametric modeling. It is standard practice to stratify the population according to observable profiles a^h of household attributes (such as family size, age, etc.). One then assumes that for a given profile $a^h = a$ the corresponding subpopulation is homogeneous in the sense that remaining variation in consumption expenditure (e.g. due to heterogeneity in individual preferences) can be described by an additive error term ϵ^h. In our notation such an approach postulates a mapping $v^h \to (p, a^h, \epsilon^h)$ and a resulting parametric model may be written in the form $c^h = f(x^h, v^h) = g(p, x^h, a^h; \theta) + \epsilon^h$.[1] Here, g is a known model function, while θ is an unknown vector of coefficients which has to be estimated from the data. Based on estimates $\hat{\theta}$, one may then compute approximations of individual elasticities.

There is an extensive literature on estimating elasticities based on such parametric approaches. Some of these estimates correspond to β_{mean}, some to β_{agg}, while others are the elasticities of g evaluated at mean or median income. For example, Houthakker (1957) assumes a double logarithmic model. In this case the covariance term is zero and hence, β_{mean} equals β_{agg}. Banks et al. (1997) rely on a more general specification (QUAIDS) and their concept of income elasticity seems to correspond to β_{agg}, since they compute an expenditure weighted mean of individual elasticities. Blundell et al. (1993) also use a QUAIDS approach and determine budget elasticities computed at the average shares and household attributes. For further empirical studies, see, e.g., Jorgenson et al (1982) or Lewbel (1989). Recently nonparametric techniques have been used to estimate aggregate elasticities β_{agg} by Chakrabarty et al. (2006).

Another branch of literature deals with the estimation of income elasticities of consumption expenditure using time-series data. In such models aggregate consumption expenditure is assumed to be a function of aggregate current and past income and other explanatory variables. Hence, under the implicit assumption that changes over time in the income distribution are captured by changes in mean income one is able to estimate an aggregate income elasticity. Important contributions in this context are, for example, Davidson et al. (1978) and Campbell and Mankiw (1990).

In this chapter we show that under general, qualitative conditions the various concepts of elasticities developed in the above sketched theoretical framework can be

[1] Models may also be formulated with respect to log consumption expenditure, $\log c^h = g(p, x^h, a^h; \theta) + \epsilon^h$, or budget shares $w^h = c^h/x^h$, $w^h = g(p, x^h, a^h; \theta) + \epsilon^h$.

identified and estimated from cross-section data (Section 3.2). In Sections 3.3-3.5 we then present an empirical study based on the microdata from the British Family Expenditure Survey (FES) (1974-1993). We estimate β_{mean}, β_{agg}, as well as close approximations to individual elasticities and the covariance term. No functional form specification of the behavioral relations is required, and no restrictive distributional assumptions have to be made. We use recent identification results described in Hoderlein and Mammen (2007) as well as nonparametric techniques for estimating regression and quantile functions as proposed by Li and Racine (2004, 2006).

We emphasize that we analyze elasticities with respect to disposable income which can be considered as an exogeneous variable. Many estimates in the literature are elasticities with respect to total expenditure. This situation is also covered by our methodology provided that total expenditure can be considered as an exogeneous variable. Otherwise, more involved nonparametric instrumental variable techniques may be applied. However, one can show that under additional assumptions our methodology offers an easy way to circumvent this problem.

The results of our empirical study are presented in Section 3.5. In particular, it turns out that the aggregate elasticity can be very different from the mean of individual elasticities. The magnitude of this difference varies from commodity to commodity. For expenditure on 'food' and 'services', as well as for 'total expenditure', aggregate income elasticity is significantly greater than the mean individual elasticities for almost all sample years. In the extreme, for expenditure on 'services' the difference can be as large as 30% of the aggregate elasticity. On the other hand, for the commodity groups 'clothing and footware' and 'fuel and light' aggregate and mean individual elasticities are quite close.

The outline of the chapter is as follows: Section 3.2 provides precise descriptions of our theoretical setup and of corresponding identification results. Section 3.3 contains information about the FES data set used in our analysis, while nonparametric estimation procedures are discussed in Section 3.4. Empirical results and conclusions are presented in Section 3.5.

3.2 Individual and Aggregate Income Elasticities

In this section we first define individual and aggregate elasticities for a population of households. The relation between these concepts is investigated. We then study the question in how far the quantities of interest can be identified and, therefore, are estimable from cross-section data. Our setup is based on general, qualitative conditions and avoids restrictive parametric model assumptions.

As explained in the introduction, a specific household with income x and a vector of further explanatory variables v determines his consumption $c \in \mathbb{R}$ of a specific commodity (e.g. food consumption) by

$$c = f(x, v),$$

The function f is assumed to be twice continuously differentiable in x. The individual income elasticity is then given by

$$\beta(x, v) := \frac{x}{c}\partial_x f(x, v) = \partial_y \log f(e^y, p), \tag{3.1}$$

where $y = \log x$.

In a large population heterogeneity in explanatory variables will generate a *joint distribution* of (x, v). Let C, X, V be corresponding generic random variables describing consumption expenditure, income and other explanatory variables of a randomly drawn household. We will use $\nu_{x,v}$ to denote the joint probability distribution of (X, V). $\nu_{x,v}$ then induces corresponding distributions of consumption $C = f(X, V)$ and individual elasticities $\beta(X, V) = \frac{X}{C}\partial_x f(X, V)$ in the population. The mean individual elasticity over the population is then given by

$$\beta_{mean} := \mathbb{E}(\beta(X, V)) = \mathbb{E}(\frac{X}{C}\partial_x f(X, V)) = \int \frac{x}{f(x, v)}\partial_x f(x, v) d\nu_{x,v} \tag{3.2}$$

Let $C_{mean} = \int f(x, v)\ d\nu_{x,v}$ denote mean consumption, while $X_{mean} = \int x\ d\nu_{x,v}$ denotes mean income. The idea motivating our definition of an aggregate income elasticity may be expressed as follows: quantify the proportional change of mean consumption in dependence of the proportion $\frac{\mu}{X_{mean}}$, when mean income is changed from X_{mean} to $\mu \neq X_{mean}$. For fixed distribution of (X, V), one considers the effect of a transformation $X \to (\frac{\mu}{X_{mean}} X)$. Obviously, $\mathbb{E}(\frac{\mu}{X_{mean}} X) = \mu$, and resulting mean consumption is given by $\mathbb{E}\left(f(\frac{\mu}{X_{mean}} X, V)\right)$. The aggregate income elasticity is then defined by

$$\begin{aligned}\beta_{agg} &= \frac{X_{mean}}{C_{mean}} \partial_\mu \mathbb{E}\left(f(\frac{\mu}{X_{mean}} X, V)\right)\Big|_{\mu = X_{mean}} = \frac{X_{mean}}{C_{mean}} \partial_\mu \int f(\frac{\mu}{X_{mean}} x, v) d\nu_{x,v}\Big|_{\mu = X_{mean}} \\ &= \frac{1}{C_{mean}} \int x \partial_x f(x, v) d\nu_{x,v} \qquad (3.3)\end{aligned}$$

It is now immediately seen that generally β_{agg} does not coincide with β_{mean}. Obviously,

$$\begin{aligned}\beta_{agg} &= \frac{1}{C_{mean}} \int f(x, v) \frac{x}{f(x, v)} \partial_x f(x, v)\ d\nu_{x,v} = \mathbb{E}(\frac{1}{C_{mean}} C \beta(X, V)) \\ &= \frac{1}{C_{mean}} C_{mean} \mathbb{E}(\beta(X, V)) + \frac{1}{C_{mean}} \mathrm{Cov}(C, \beta(X, V)) \\ &= \beta_{mean} + \frac{1}{C_{mean}} \mathrm{Cov}(C, \beta(X, V)) \qquad (3.4)\end{aligned}$$

Let us now consider the question which of the above quantities are identifiable from cross-section data. A basic problem is that individual preferences and, hence, the parameters v are not directly observable. However, all expenditure surveys provide information about important household attribute profiles a, as for example household size, employment status, age of household members, etc. Let us thus analyze the situation that there is an i.i.d. sample (C_i, X_i, A_i), $i = 1, \ldots, n$, containing information about consumption, income and household attributes of n randomly selected households. Following the above notation, the distribution of (C_i, X_i, A_i) corresponds to the distribution of generic variables (C, X, A). The introduction of attribute profiles is crucial, since generally A will be correlated with the unobservable random variable V. Let $\nu_{x,a}$ denote the joint distribution of (X, A), while $\nu_{v|x,a}$ stands for the conditional distribution of V given (X, A). Note that in our setup X denotes disposable income of a household (and not total expenditure). X (and $Y = \log X$) are thus assumed to be exogeneous variables. A standard assumption which implicitly or explicitly provides the very basis for almost all theoretical and applied work to be found in the literature is as follows:[2] If attribute profiles a provide sufficient information about the characteristics of an household, then for the subgroups of all household with the same attributes $A = a$ the variation in V only reflects variation in individual preferences which may be assumed to be independent of income. More formally, further analysis will rest upon the following assumption:

> **Assumption** (*Conditional independence of X and V given A*):
> For every income level x, $\nu_{v|x,a} = \nu_{v|a}$, where $\nu_{v|a}$ denotes the conditional distribution of V given $A = a$.

Let us now first study identification of β_{mean} and β_{agg}. Set $Y = \log X$,

$$\bar{c}(y, a) := \mathbb{E}(C \mid Y = y, A = a), \quad \bar{c}_{log}(y, a) := \mathbb{E}(\log C \mid Y = y, A = a),$$

and let $\nu_{y,a}$ be the joint distribution of (Y, A). Under the above assumption,

$$\begin{aligned}\beta_{mean} &= \mathbb{E}(\beta(X, V)) = \mathbb{E}(\frac{X}{C}\partial_x f(X, V)) = \int \left(\int \frac{x}{f(x, v)} \partial_x f(x, v) d\nu_{v|a} \right) d\nu_{x,a} \\ &= \int \partial_y \left(\int \log f(e^y, v) d\nu_{v|a} \right) d\nu_{y,a} = \int \partial_y \bar{c}_{log}(y, a) d\nu_{y,a} \qquad (3.5)\end{aligned}$$

and

$$\begin{aligned}\beta_{agg} &= \frac{1}{C_{mean}} \int x \partial_x f(x, v) d\nu_{x,v} = \frac{1}{C_{mean}} \int \partial_y \left(\int f(e^y, v) d\nu_{v|a} \right) d\nu_{y,a} \\ &= \frac{1}{C_{mean}} \int \partial_y \bar{c}(y, a) d\nu_{y,a} \qquad (3.6)\end{aligned}$$

The functions $\bar{c}(y, a)$ and $\bar{c}_{log}(y, a)$ are well identified regression functions. Nonparametric regression procedures can be used to determine estimates $\widehat{\bar{c}(y, a)}$ and $\widehat{\bar{c}_{log}(y, a)}$ by regressing C_i on (Y_i, A_i) and $\log C_i$ on (Y_i, A_i), respectively. By (3.5) and (3.6) the

---

[2] Parametric models of demand may be written in the form $C_i = g(p, X_i, A_i; \theta) + s(p, X_i, A_i; \theta) \cdot \epsilon_i$ (or $\log C_i = g(p, X_i, A_i; \theta) + s(p, X_i, A_i; \theta) \cdot \epsilon_i$) for some prespecified functions g and s, where $\epsilon_1, \epsilon_2, \ldots$ are i.i.d random errors ϵ_i with $\mathbb{E}(\epsilon_i) = 0$ and $Var(\epsilon_i) = 1$ which are assumed to be *independent* of X_i, A_i. The function s may be used to account for possible heteroscedasticity, while θ denotes some unknown vector of parameters that have to be estimated from the data. In such a setup the "error term" $\epsilon_i \equiv \epsilon(V_i)$ obviously captures remaining heterogeneity of C_i for given (X_i, A_i). Conditional independence of X_i and V_i given A_i is an immediate consequence.

elasticities β_{mean} and β_{agg} then are (suitably scaled) **average derivatives** of $\bar{c}_{log}(y, a)$ and $\bar{c}(y, a)$, which may be estimated by $\frac{1}{n}\sum_{i=1}^{n}\partial_y\widehat{\bar{c}_{log}(Y_i, A_i)}$ and $\frac{1}{\bar{C}n}\sum_{i=1}^{n}\partial_y\widehat{\bar{c}(Y_i, A_i)}$, respectively, where $\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i$. Details of our estimation procedures are given in Section 3.4.

Identification of individual elasticities is, of course, a much more difficult problem. Quite surprisingly, in a general setup it is possible to get some "close" approximations. Our identification strategy is based on the approach of Hoderlein and Mammen (2007).

For $0 \leq \tau \leq 1$, let $k(\tau; y, a)$ denote the conditional τ-quantile of $\log C$ given $Y = y$ and $A = a$. More formally, $P(\log C \leqslant k(\tau; y, a)|Y = y, A = a) = \tau$. We will assume that $k(\tau; y, a)$ is continuously differentiable with respect to y and that $k(\tau; y, a)$ is strictly increasing in τ for all (y, a). For any given (c, y, a) there then exists some $\tau_{c,y,a}$ such that $\log c = k(\tau_{c,y,a}; y, a)$ Under some mild regularity conditions on the distribution $\nu_{y,a}$, the results of Hoderlein and Mammen (2007) then imply that

$$\beta_{c,y,a} := \mathbb{E}\big[\partial_y \log f(e^y, V, p)|Y = y, A = a, C = c\big] = \partial_z k(\tau_{c,y,a}; z, a)\Big|_{z=y}. \quad (3.7)$$

We will refer to $\beta_{c,y,a}$ as a "local" elasticity. By definition,

$$\beta_{c,y,a} = \mathbb{E}\left(\beta(X, V)\ |Y = y, A = a, C = c\right),$$

and thus it is the conditional mean of $\beta(X, V)$ over a subpopulation of households with log income y, attributes a *and* consumption expenditure equal to c. Although households within such a subpopulation can still be heterogeneous in v, they show the same consumption behavior given y and a. Consequently, when using i.i.d. data providing information about consumption C_i , income X_i and household attributes A_i, $i = 1, \ldots, n$, approximating the individual elasticity $\beta(X_i, V_i)$ by the conditional mean $\beta_i := \beta_{C_i,Y_i,A_i} = \mathbb{E}\big[\partial_y \log f(e^{Y_i}, V, p)|Y = Y_i, A = A_i, C = C_i\big], \quad i = 1, \ldots, n$ is the best we can do on the basis of the available information. Local and individual income elasticities will coincide if, for given (y, a), there is a one-to-one relation between consumption c and (preference) parameters v.[3] Identification of demand function under such "monotonicity" constraints has been considered by Matzkin (2003).

By relying on nonparametric quantile estimation techniques, nonparametric estimates $\widehat{k(\tau; y, a)}$ of conditional quantile functions and their derivatives can be determined from cross-section data. For any observation (C_i, Y_i, A_i) the corresponding conditional quantile position of $\log C_i$ given (Y_i, A_i) can be computed in a straightforward way which then leads to estimates $\widehat{\beta}_i$ of local elasticities.

Local elasticities provide a mean to estimate $\text{Cov}(C, \beta(X, V))$. Obviously, $\mathbb{E}(\beta_i) = \mathbb{E}(\beta_{C,Y,A}) = E(\beta(X, V)) = \beta_{mean}$, and

$$\begin{aligned}\text{Cov}(C, \beta(X, V)) &= \mathbb{E}(C\beta(X, V)) - C_{mean}\mathbb{E}(\beta(X, V)) \\ &= \mathbb{E}\left[C\mathbb{E}(\beta(X, V)|Y, A, C)\right] - C_{mean}\beta_{mean} \\ &= \mathbb{E}\left[C\beta_{C,Y,A}\right] - C_{mean}\beta_{mean} = \text{Cov}(C, \beta_{C,Y,A}) = \text{Cov}(C_i, \beta_i) \quad (3.8)\end{aligned}$$

Remark: As mentioned in the introduction many contributions in the literature aim at estimating elasticities with respect to total expenditure on nondurables, "budget" for short, which is usually considered as an endogeneous variable. The behavioral relations imply that $C_{tot} = f_{tot}(X, V)$ and $C = f(X, V)$, where C_{tot} denotes budget and C refers to consumption expenditure on another commodity. Assuming monotonicity of f_{tot} in income, C can be rewritten as a function of budget, $C = \tilde{f}(C_{tot}, V) = \tilde{f}(f_{tot}(X, V), V)$. The elasticity with respect to budget is then given by $\tilde{\beta}(C_{tot}, V) = \frac{C_{tot}}{C}\partial_{c_{tot}}\tilde{f}(C_{tot}, V)$. Taking derivatives yields

$$\frac{C}{X}\beta(X, V) = \frac{C}{C_{tot}}\tilde{\beta}(C_{tot}, V) \cdot \frac{C_{tot}}{X}\beta_{tot}(X, V), \quad (3.9)$$

where β_{tot} denotes the elasticity of C_{tot} with respect to income. Under the additional

[3] Such an assumption is made in any parametric model of demand. If, for example,

$$C_i = g(p, X_i, A_i; \theta) + s(p, X_i, A_i; \theta) \cdot \epsilon_i$$

for some prespecified g, s, then

$$\beta_i = \mathbb{E}\big[\partial_y \log f(e^{Y_i}, V)|Y = Y_i, A = A_i, C = C_i\big] = \partial_y \log(g(p, e^{Y_i}, A_i; \theta) + s(p, e^{Y_i}, A_i; \theta) \cdot \epsilon_i)$$

equals the individual elasticity $\beta(X_i, V_i)$ (recall that $\epsilon_i \equiv \epsilon(V_i)$).

assumption that local elasticities are equal to individual elasticities, we can infer from (3.9) that the mean elasticity $\beta_{mean,tot} := \mathbb{E}(\tilde{\beta}(C_{tot}, V))$ with respect to budget corresponds to

$$\beta_{mean,tot} = \mathbb{E}\left(\frac{\beta(X,V)}{\beta_{tot}(X,V)}\right) = \mathbb{E}\left(\frac{\beta_{C,Y,A}}{\beta_{tot;C_{tot},Y,A}}\right), \tag{3.10}$$

where $\beta_{tot;C_{tot},Y,A}$ denotes local elasticities of C_{tot} with respect to income. This elasticity may then be estimated by $\widehat{\beta}_{mean,tot} = \frac{1}{n}\sum_i \frac{\hat{\beta}_i}{\hat{\beta}_{tot,i}}$. Similar arguments may then be used to show that the corresponding aggregate elasticity $\beta_{agg,tot}$ can also be determined from local elasticities.

3.3 Data Description

Our empirical analysis bases on the British Family Expenditure Survey, which contains cross-section data on consumption expenditure, income and socioeconomic characteristics of British households. FES was launched in the late 50s but due to changes in survey design and the following inconsistency in variable definitions we restrict our analysis to the period 1974-1993.

Annually, FES asks approximately 7000 households to keep a detailed account of their expenditures on a variety of commodity groups for 14 consecutive days. Depending on how necessary the good is, one might expect different demand behavior for different categories of goods. Therefore we perform our analysis for the major four commodity groups: 'food', 'fuel and light', 'services', and 'clothing and footware,' as well as for total (nondurable) expenditure. As far as the income variable is concerned, it is the natural logarithm of the disposable non-property income, which is obtained by deducting investment income and all taxes from total income.[4]

A correct specification of the econometric model and the need for stratification of the population on attributes induced by the theoretical model motivates the inclusion

[4]Following HBAI standards, household incomes are obtained by extracting relevant items from the elementary database. The task of elaborating the database and specifying consistent variables has mainly been accomplished by Jürgen Arns and described in Arns (2006) and Arns and Bhattacharya (2005). His careful work is gratefully acknowledged.

of further explanatory variables for household demand in the empirical analysis. The following variables[5] have been chosen in our analysis: number of adults, children and persons working in the household, and age as well as employment status of the household's head.

3.4 Estimation Procedures

In this section we give a detailed description of our procedure for estimating the quantities β_{agg}, β_{mean}, β_i, and $\mathrm{Cov}(C, \beta(X, V))$ from cross-section data. It is assumed that for a given time period of interest there are observations (C_i, X_i, A_i) of consumption, income and attributes, $i = 1, \ldots, n$, for an i.i.d. sample of n households.

3.4.1 Estimation of β_{agg} and β_{mean}

By definition in (3.6) and (3.5) β_{agg} and β_{mean} are average derivatives with respect to the regression functions $\bar{c}_{(}y, a) = \mathbb{E}(C|Y = y, A = a)$ and $\bar{c}_{\log}(y, a) = \mathbb{E}(\log C_i|Y = y, A = a)$, where $Y = \log X$. In other words, in order to estimate β_{agg} we have to regress C on (Y, A), while for approximating β_{mean} one has to regress $\log C$ on (Y, A).

In principle, estimation could be based on valid parametric models for $\bar{c}(y, a)$ and $\bar{c}_{\log}(y, a)$, respectively. A straightforward approach would be to use a model for $\bar{c}(y, a)$, which is quadratic in log income and age, linear in the number of adults and children and dummies for employment status and allows for interaction between log income and age. Such a model could be then estimated by the least squares method and the derivative $\partial_y \bar{c}(y, a)$ could be computed as a function of estimated parameters. However, as shown by Chakrabarty et al. (2006) for the FES data, such a model suffers from misspecification according to the Ramsey (1969) RESET test for all commodity groups but for 'fuel and light' and 'clothing and footware' for some years.

In this chapter we therefore adopt a nonparametric approach for estimating $\bar{c}$ and $\bar{c}_{\log}$. We rely on the methodology developed in Li and Racine (2004) which is well-

[5] For a more detailed exposition of the elementary data set and variable definitions we refer to the FES Handbook by Kemsley et al. (1980).

adapted to the fact that some regressors are continuous (log income and age), while others are categorical (number of adults and children, employment status).

More precisely, the vector A_i of household attributes is split into the continuous attribute 'age of household head', denoted by age_i, and a vector $\tilde{A}_i$ of six discrete variables: three dummies for the employment status of the household's head (unemployed/unoccupied, self-employed, and retired), number of persons in the household, number of children, and number of persons working. Among the discrete variables we distinguish unordered and ordered ones. We treat the three employment status dummies as unordered and the remaining discrete regressors as ordered variables. Following Li and Racine (2004), estimates of the regression functions are obtained by local linear weighted regressions.

Let $Z_i = (Y_i, \text{age}_i, \tilde{A}_i^T)^T$, $i = 1, \ldots, n$, denote the individual vectors of all explanatory variables. For a given point $z = (y, \text{age}, \tilde{a}^T)^T$, estimates of $\bar{c}(y, a)$ and $\partial_y \bar{c}(y, a)$ are then determined by $\widehat{\bar{c}(y, a)} := \hat{\zeta}_0$ and $\widehat{\partial_y \bar{c}(y, a)} := \hat{\zeta}_1$, respectively, where $\hat{\zeta}_0$, $\hat{\zeta}_1$, $\hat{\zeta}_2$ minimize

$$\sum_{i=1}^{n} \left[C_i - \zeta_0 - \zeta_1(Y_i - y) - \zeta_2(\text{age}_i - \text{age})\right]^2 W_{i,\mathbf{h}}(z),$$

over all possible values $\zeta_0, \zeta_1, \zeta_2$. Hereby, $W_{i,\mathbf{h}}$ is a kernel weight for household i at point z, which depends on the bandwidth vector $\mathbf{h}$. These weights are computed as a product of univariate kernel functions, where the functional forms of the kernels are chosen according to the nature of the respective variables: Epanechnikov kernels $\kappa(\cdot)$ for continuous, Aitchison and Aitken (1976) kernel $l^u(\cdot)$ for unordered categorical, and Wang and Van Ryzin (1981) kernel $l^o(\cdot)$ for ordered discrete variables.[6] Consistency

[6]More precisely,

$$W_{i,\mathbf{h}}(z) = \kappa\Big(\frac{Y_i - y}{\mathbf{h}_1}\Big)\kappa\Big(\frac{age_i - age}{\mathbf{h}_2}\Big) \prod_{s=1}^{3} l^u(\tilde{A}_{is}, \tilde{a}_s, \mathbf{h}_s) \prod_{s=4}^{6} l^o(\tilde{A}_{is}, \tilde{a}_s, \mathbf{h}_s),$$

where κ, l^u, and l^o are continuous, unordered discrete, and ordered discrete kernels, respectively. They are defined by $\kappa(u) = \begin{cases} \frac{3}{4\sqrt{5}}\big(1 - \frac{1}{5}u^2\big), & \text{if } u^2 < 5 \\ 0, & \text{else} \end{cases}$, $l^u(\tilde{A}_{is}, \tilde{a}_s, \mathbf{h}_s) = \begin{cases} 1 - \mathbf{h}_s, & \text{if } \tilde{A}_{is} = \tilde{a}_{as} \\ \mathbf{h}_s/(o_s - 1), & \text{else} \end{cases}$, and $l^o(\tilde{A}_{is}, \tilde{a}_s, \mathbf{h}_s) = \begin{cases} 1 - \mathbf{h}_s, & \text{if } \tilde{A}_{is} = \tilde{a}_s \\ \frac{1}{2}(1 - \mathbf{h}_s)\mathbf{h}_s^{|\tilde{A}_{is} - \tilde{a}_s|}, & \text{else} \end{cases}$, where o_s is the number of possible outcomes of $\tilde{A}_{is}$.

and asymptotic normality of this estimators follow from the results of Li and Racine (2004).

Similarly, estimates $\widehat{\bar{c}_{\log}(y,a)} := \hat{\zeta}_0^*$ and $\partial_y \widehat{\bar{c}_{\log}(y,a)} := \hat{\zeta}_1^*$ are calculated from the minimizers $\hat{\zeta}_0^*$, $\hat{\zeta}_1^*$, $\hat{\zeta}_2^*$ of

$$\sum_{i=1}^{n} \big[\log C_i - \zeta_0^* - \zeta_1^*(Y_i - y) - \zeta_2^*(\text{age}_i - \text{age}) \big]^2 W_{i,\mathbf{h}}(z).$$

By (3.6) and (3.5) this then leads to the estimates

$$\hat{\beta}_{agg} = \frac{1}{\bar{C}n} \sum_{i=1}^{n} \partial_y \widehat{\bar{c}(Y_i, A_i)}, \quad \hat{\beta}_{mean} = \frac{1}{n} \sum_{i=1}^{n} \partial_y \widehat{\bar{c}_{log}(Y_i, A_i)}.$$

The optimal smoothing parameters for estimating $\bar{c}$ and $\bar{c}_{log}$, which are denoted by $\mathbf{h}_{CV}$ and $\mathbf{h}_{CV}^*$, respectively, are chosen by a least-squares cross-validation algorithm as described in Racine and Li (2004).[7] However, recall that we are interested in estimating the corresponding average derivative of the regression function and not the regression function itself. Averaging reduces variability of the estimate but not its bias. In a similar context, Härdle and Stoker (1989) show that by applying 'undersmoothing' bandwidths parametric rates of convergence can be achieved for average derivative estimators. We therefore determine $\hat{\beta}_{agg}$ and $\hat{\beta}_{mean}$ by using bandwidths $0.8\mathbf{h}_{CV,1}$ and $0.8\mathbf{h}_{CV,1}^*$ for log income, respectively.[8] Additionally, for the sake of stability of results, while computing the average derivative and the covariance term we neglect the highest and the lowest 0.5% of the values of the point derivatives.

Separately for each period, standard errors of $\hat{\beta}_{agg}$ and $\hat{\beta}_{mean}$ can be obtained by bootstrap. For i.i.d bootstrap resamples $(\log C_1^*, Z_1^*), \ldots, (\log C_n^*, Z_n^*)$ the distributions of $\hat{\beta}_{agg} - \beta_{agg}$, $\hat{\beta}_{mean} - \beta_{mean}$ are approximated by the conditional distribution of $\hat{\beta}_{agg}^* - \hat{\beta}_{agg}$, $\hat{\beta}_{mean}^* - \hat{\beta}_{mean}$ given (C_i, Y_i, A_i), $i = 1, \ldots, n$. Theoretical support for

[7]Numerical search for optimal smoothing parameters was performed using the N library made available by Jeff Racine. The estimation procedure itself was programmed in MATLAB and the corresponding routines are available from authors upon request.

[8]The estimates of $\hat{\beta}_{agg}$ and $\hat{\beta}_{mean}$ obtained with this bandwidth vector were very similar to those obtained when using factors 0.7 or 0.9. This may indicate that we are close to the optimal bandwidth (for the average derivative estimator).

the use of such a naive bootstrap in the context of average derivative estimation can be found in Härdle and Hart (1991).

3.4.2 Estimation of Local Elasticities and Cov$(C, \beta(X, V))/C_{mean}$

As already explained in Section 3.2, the strategy for estimating the individual values of the local elasticities β_i, $i = 1, \ldots, n$, stems from Hoderlein and Mammen (2007). We apply a two-step procedure. In the first step, we determine estimates $\hat{\tau}_i$ of the quantiles $\tau_i := \tau_{C_i,Y_i,A_i}$ with $\log C_i = k(\tau_{C_i,Y_i,A_i}; Y_i, A_i)$, i.e., of the quantile positions of $\log C_i$ in the distribution of log expenditure across the subpopulation with log income and attributes equal to (Y_i, A_i). In the second step, one estimates the partial derivative of $k(\tau_{C_i,Y_i,A_i}; y, A_i)$ at $y = Y_i$. As we do not want to impose any restrictive assumptions on the shape of the conditional quantile function $k(\cdot)$, our approach again relies on nonparametric procedures. As in the case of estimation of $\bar{c}(y, a)$ and $\bar{c}_{\log}(y, a)$ described above, we have to account for the presence of both discrete and continuous variables. We therefore apply a general method for quantile estimation which has been developed in a recent work by Li and Racine (2006).

Consistent estimators of τ_i, $i = 1, \ldots, n$, are then given by

$$\hat{\tau}_i = \frac{\sum_{j=1}^{n} G\big(\frac{\log C_j - \log C_i}{\mathbf{h}_0}\big) W_{j,\mathbf{h}}(Z_i)}{\sum_{j=1}^{n} W_{j,\mathbf{h}}(Z_i)},$$

where G is the cumulative continuous kernel function, i.e., $G(t) = \int_{-\infty}^{t} \kappa(u)du$, $\mathbf{h}_0$ is the bandwidth parameter for C, and $W_{j,\mathbf{h}}(Z_i)$ is the kernel weight for the household j at Z_i, which has already been defined above. Bandwidth parameters for the estimation of τ_i were chosen through a numerical search algorithm presented for conditional density estimation in Hall et al. (2004) method and properly adjusted for the estimation of conditional cumulative distribution functions as advocated by Li and Racine (2006).

In the second step, we perform a local linear[9] quantile regression at the quantile $\hat{\tau}_i$.

[9]From the theoretical point of view local quadratic smoother outperforms the linear one in estimating the derivative of $k(\tau; y, a)$. However, in our application local quadratic regression (even for large bandwidths) leads to more instable estimates. In particular, for food expenditure we then

More precisely, for each $i = 1, \dots, n$ we calculate the values $\hat{\eta}_{0,i}, \hat{\eta}_{1,i}, \hat{\eta}_{2,i}$ minimizing

$$\sum_{j=1}^{n} \rho_{\hat{\tau}_i} \big[\log C_j - \eta_{0,i} - \eta_{1,i}(Y_j - Y_i) - \eta_{2,i}(\text{age}_j - \text{age}_i) \big] W_{j,\mathbf{h}}(Z_i), \tag{3.11}$$

with respect to all $\eta_{0,i}$, $\eta_{1,i}$, and $\eta_{2,i}$, where $\rho_\tau(u) = u[\tau - I(u \le 0)]$ is a 'check function' typical for quantile regression problems.[10] In the above regression $\hat{\eta}_{1,i}$ estimates the partial derivative of $k(\tau_{C_i,Y_i,A_i}; y, A_i)$ at $y = Y_i$, and therefore $\hat{\beta}_i := \hat{\eta}_{1,i}$. By (3.8) these estimates $\hat{\beta}_i$ of local elasticities can then be used to estimate $\text{Cov}(C, \beta(X, V))$:

$$\text{Cov}(\widehat{C, \beta(X}, V)) = \frac{1}{n} \sum_{i=1}^{n} (C_i - \bar{C})(\hat{\beta}_i - \bar{\beta}),$$

where $\bar{\beta} = \frac{1}{n} \sum_{i=1}^{n} \hat{\beta}_i$. Whereas estimation of average derivatives (as in the case of β_{agg} or β_{mean}) relies on a smaller bandwidth for log income than the optimal one for estimating the regression function, point derivatives of quantiles should be estimated using a *larger* bandwidth. Since the direct data-driven bandwidth selection methods in this situation are still an open question, we proceed as follows. First, we multiply the cross-validated bandwidth for log income by a factor 1.5 which results in $\mathbf{h}_d^*$ with $\mathbf{h}_{d,1}^* = 1.5\mathbf{h}_{CV,1}^*$ and $\mathbf{h}_{d,s}^* = \mathbf{h}_{CV,s}^*$, for $s > 1$.[11] Then, as advocated by Yu and Jones (1998), in order to obtain a suitable bandwidth for quantile derivative estimation we adjust smoothing parameters for log income $\mathbf{h}_{d,1}^*$ and age $\mathbf{h}_{d,2}^*$ in dependence of $\hat{\tau}_i$ by multiplying them by a factor $\left[\frac{\hat{\tau}_i(1-\hat{\tau}_i)}{\phi[\Phi^{-1}(\hat{\tau}_i)]^2}\right]^{1/6}$. Here ϕ and Φ are the pdf and the cdf of the standard normal distribution, respectively. For discrete variables we use the same bandwidths as in the mean regression case.[12]

---

obtain an implausibly high percentage of negative elasticities.

[10]It is important to note that the presence of this function is the only difference between a typical (mean) regression and a quantile regression.

[11]Our estimates of β and $\text{Cov}(C, \beta(X, V))$ are stable with respect to changes in this multiplier between 1 and 2.

[12]Numerical search for optimal bandwidths in the first step was carried out by the N library made available by Jeff Racine. Estimators for both estimation steps were programmed in MATLAB. The solution to (3.11) was found by the interior point (Frisch-Newton) algorithm implemented in the RQ.m routine and described by Portnoy and Koenker (1997). Program codes for these routines are available from authors upon request.

3.4.3 Inference about $\mathbf{Cov}(C, \beta(X, V))/C_{mean}$

Having estimated the aggregate elasticity, the mean of individual elasticities, and the covariance term, it is of interest to assess whether the difference $\beta_{agg} - \beta_{mean}$, or equivalently, the covariance term $\text{Cov}(C, \beta(X, V))/C_{mean}$ is significantly different from zero. We propose two different tests for equality of β_{agg} and β_{mean}.

In the first test, for each year of the sample we test the null hypothesis $H_0 : \text{Cov}(C, \beta(X, V))/C_{mean} = 0$. As there does not exist a closed form for the asymptotic standard error of the covariance term, in order to analyze its significance, the test is based on bootstrap confidence intervals. Bootstrap resamples $(\log C_1^*, Z_1^*), \ldots, (\log C_n^*, Z_n^*)$ are generated by drawing independently, with replacements n observations from the original sample $(\log C_1, Z_1), \ldots, (\log C_n, Z_n)$. For each bootstrap sample corresponding estimates $\widehat{\text{Cov}(C, \beta(X}, V))^*$ and $\bar{C}^* = \frac{1}{n}\sum_{i=1}^n C_i^*$ are determined. The distribution of $(\widehat{\text{Cov}(C, \beta(X}, V))/\bar{C} - \text{Cov}(C, \beta(X, V))/C_{mean})$ is then approximated by the corresponding bootstrap distribution $(\widehat{\text{Cov}(C, \beta(X}, V))^*/\bar{C}^* - \widehat{\text{Cov}(C, \beta(X}, V))/\bar{C})$. We compute 95% confidence intervals for $\text{Cov}(C, \beta(X, V))/C_{mean}$ by

$$\text{C.I.} = [\widehat{\text{Cov}(C, \beta(X}, V))/\bar{C} - t_{0.975}^*,\ \widehat{\text{Cov}(C, \beta(X}, V))/\bar{C} - t_{0.025}^*],$$

where t_α^* denotes the α quantile of the generated bootstrap distribution $(\widehat{\text{Cov}(C, \beta(X}, V))^*/\bar{C}^* - \widehat{\text{Cov}(C, \beta(X}, V))/\bar{C})$. As shown by Koenker (1994), this type of bootstrap performs very well in quantile regression problems under heteroscedasticity, which is present in our data.

In the second test, we consider the significance of the average difference between β_{agg} and β_{mean} over the sample period (1974-1993). We obtain two series $\{\hat{\beta}_{agg,t}\}$ and $\{\hat{\beta}_{mean,t}\}$ for $t = 74, \ldots, 93$ and test their equality by means of the Wilcoxon (1945) test for matched pairs. Additionally, we perform the Wilcoxon signed-rank test for zero mean of $\text{Cov}(C, \beta(X, V))/C_{mean}$ based on observations $\widehat{\text{Cov}(C, \beta(X}, V))_t/\bar{C}_t$, $t = 74, \ldots, 93$. All empirical results are given in the next section.

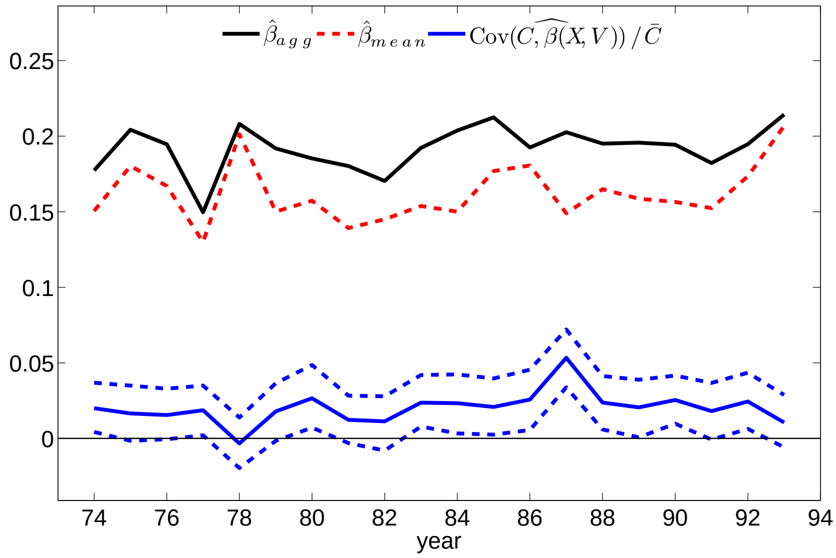

Figure 3.1: Estimates of elasticities and the covariance term for 'food expenditure'

3.5 Estimation Results and Conclusions

Tables 3.1-3.5 report our estimates of the aggregate elasticity $\hat{\beta}_{agg}$ (first column) and the mean individual elasticity $\hat{\beta}_{mean}$ (second column) for each commodity group. The third and the fourth column provide corresponding estimates $\bar{\beta} = \frac{1}{n}\sum_i \hat{\beta}_i$ of mean local elasticities and of $\mathrm{Cov}(\widehat{C, \beta(X, V)})/\bar{C}$, respectively. In the parentheses next to the estimates we report their bootstrapped standard errors. Furthermore, in Figures 3.1-3.5 we plot the time-series of estimates $\{\hat{\beta}_{agg,t}\}$, $\{\hat{\beta}_{mean,t}\}$, as well as $\mathrm{Cov}(\widehat{C, \beta(X, V)})_t/\bar{C}_t$, $t = 74, \ldots, 93$, with corresponding 95% confidence intervals.

Our estimation results lead to the following conclusions:

1) There are large differences in the magnitude of the elasticities among different commodity groups. In particular, an increase in aggregate income of 1% drives up aggregate expenditure for 'food' or 'fuel and light' by approximately 0.2%, whereas for expenditure on 'services' this increase is of roughly 1%. Total expenditure for all nondurable goods rises by about 0.5% on average.

2) From Figures 3.6-3.8, where we present kernel density estimates of the distribution of local elasticities $\hat{\beta}_i$ for different commodity groups in 1993, we see

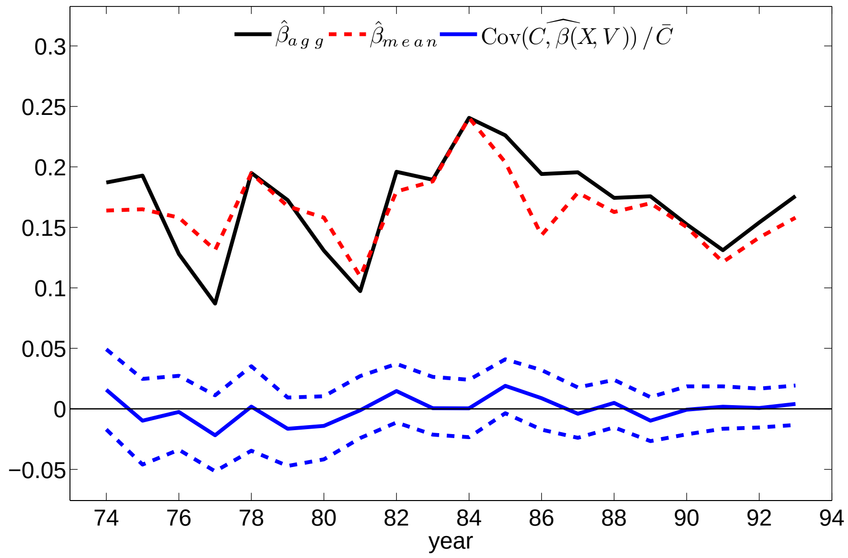

Figure 3.2: Estimates of elasticities and the covariance term for 'fuel and light expenditure'

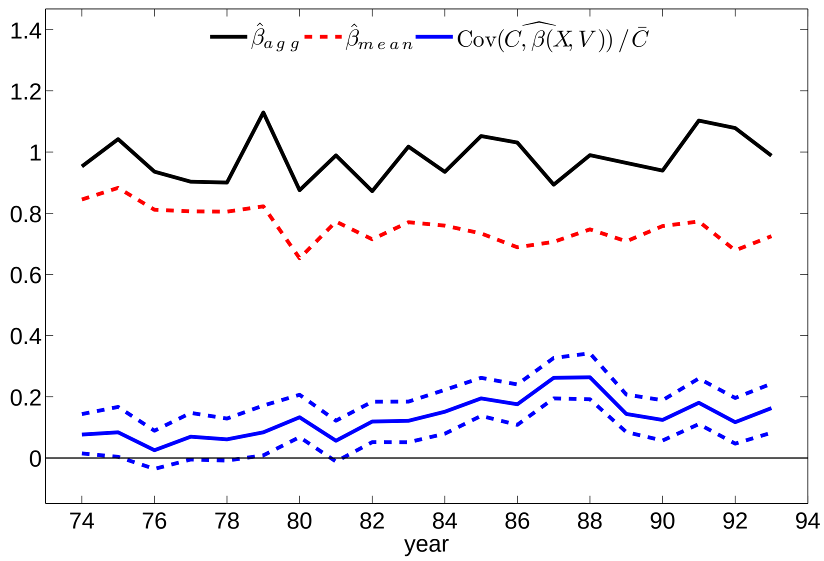

Figure 3.3: Estimates of elasticities and the covariance term for 'services expenditure'

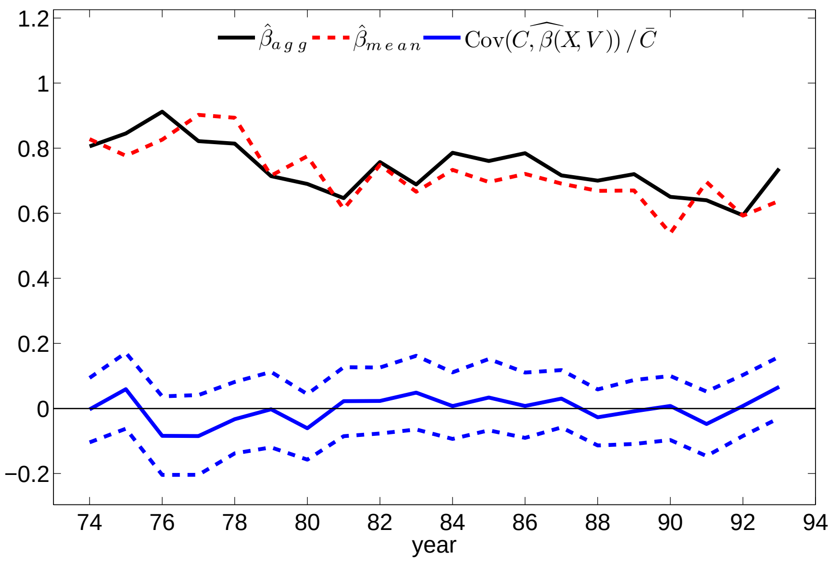

Figure 3.4: Estimates of elasticities and the covariance term for 'clothing and footware expenditure'

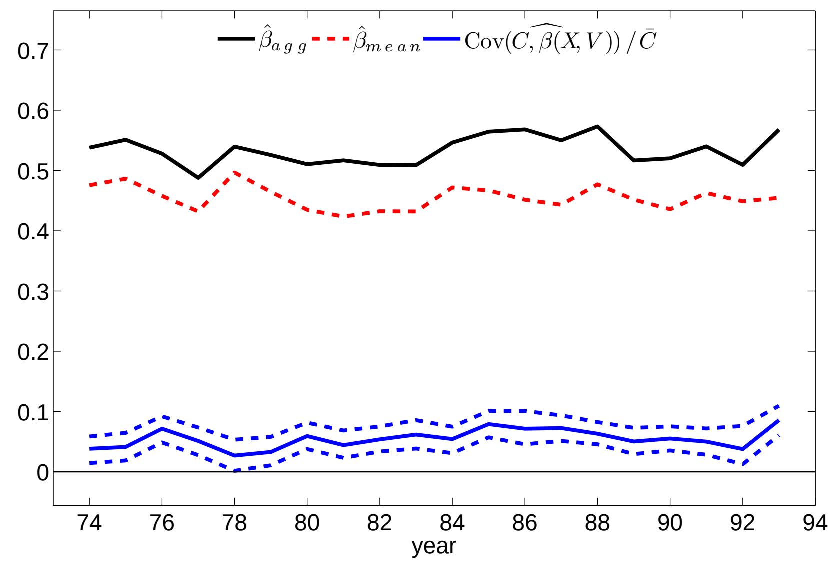

Figure 3.5: Estimates of elasticities and the covariance term for 'total (nondurable) expenditure'

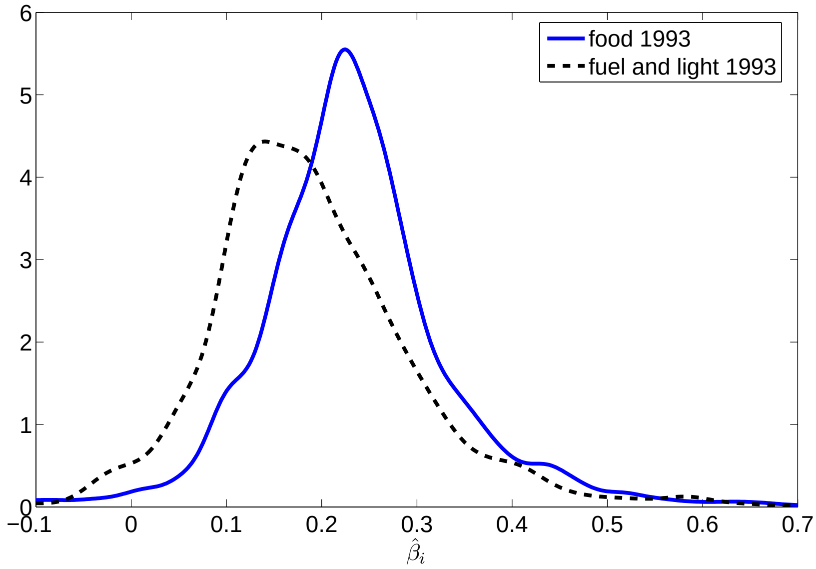

Figure 3.6: Distribution of $\hat{\beta}_i$ for expenditures on 'food' and 'fuel and light' in 1993.

that these distributions are unimodal and exhibit a significant spread. This last feature indicates a substantial degree of heterogeneity in demand behavior across the population. Furthermore, according to the Jarque-Bera test, these distributions are very far from being normal for all years and for all commodity groups.[13]

3) The estimates of elasticities seem to be fairly stable over time. During the period 1974-1993, one can observe no pronounced trend in the estimates of both the aggregate elasticity and the mean individual elasticity.

4) The estimates of $\mathbb{E}(\beta(X,V)) = \beta_{mean}$ obtained by $\hat{\beta}_{mean}$ and by the average $\bar{\beta} = \frac{1}{n}\sum_i \hat{\beta}_i$ of local elasticities are of very similar magnitude, which could serve as a support for the reliability and robustness of these estimates. Further, for most commodity groups and sample years we can recover the relationship from the proposition saying that $\beta_{agg} = \beta_{mean} + \mathrm{Cov}(C, \beta(X,V))/C_{mean}$, which

[13] Note that for a small group of households the elasticity is estimated to be negative. The size of this group varies by commodity group and year and is of magnitude of one to three percent of the sample. The explanation for the occurrence of negative elasticities is the methodical artifact of the nonparametric smoother applied in our work.

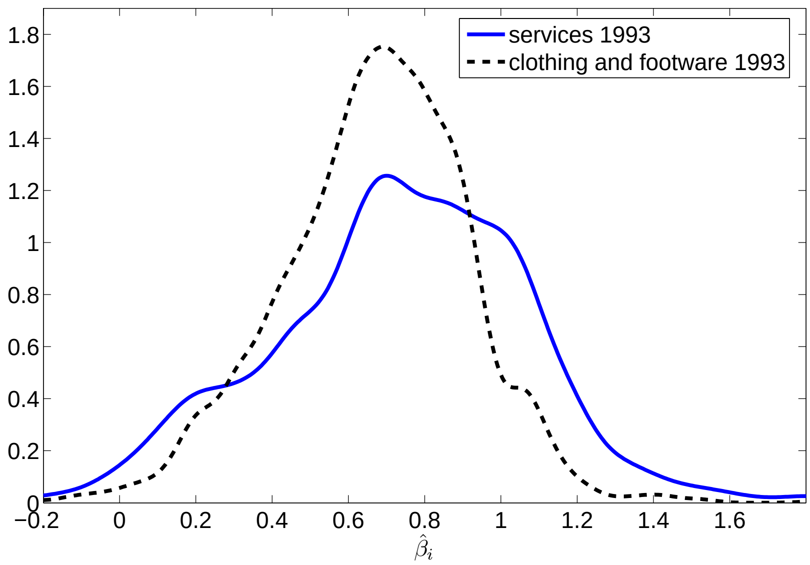

Figure 3.7: Distribution of $\hat{\beta}_i$ for expenditures on ‘services’ and ‘clothing and footware’ in 1993.

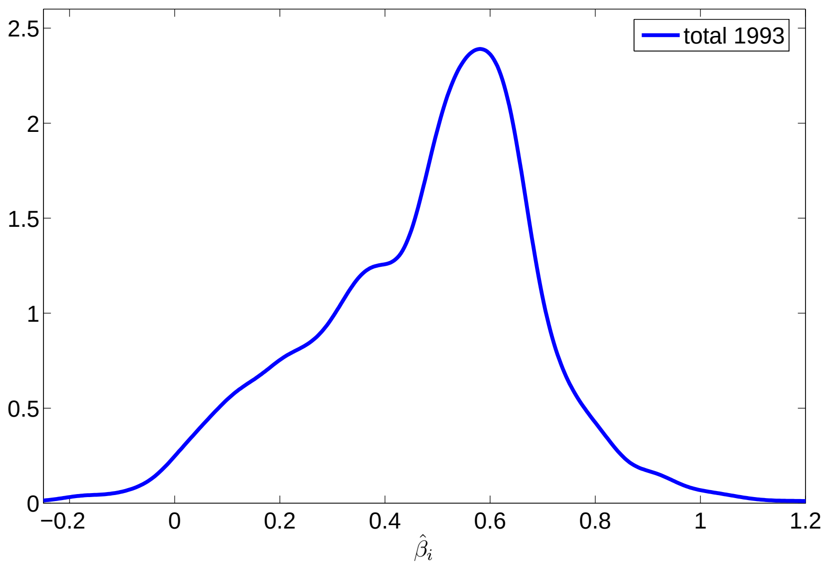

Figure 3.8: Distribution of $\hat{\beta}_i$ for expenditures on ‘total (nondurable) expenditure’ in 1993.

year	$\hat{\beta}_{agg}$	$\hat{\beta}_{mean}$	$\bar{\beta}$	$\widehat{\text{Cov}(C, \beta(X,V))}/\bar{C}$	C.I.
1974	0.177 (0.017)	0.151 (0.017)	0.148 (0.020)	0.020 (0.008)	*[0.004,0.037]
1975	0.204 (0.020)	0.180 (0.019)	0.173 (0.021)	0.017 (0.009)	[-0.002,0.035]
1976	0.195 (0.017)	0.167 (0.018)	0.161 (0.020)	0.015 (0.009)	[-0.001,0.033]
1977	0.150 (0.019)	0.130 (0.019)	0.132 (0.022)	0.019 (0.009)	*[0.002,0.035]
1978	0.208 (0.018)	0.201 (0.018)	0.194 (0.023)	-0.003 (0.009)	[-0.020,0.014]
1979	0.192 (0.021)	0.150 (0.020)	0.156 (0.022)	0.018 (0.010)	[-0.002,0.037]
1980	0.185 (0.019)	0.157 (0.021)	0.161 (0.022)	0.027 (0.010)	*[0.007,0.049]
1981	0.180 (0.017)	0.139 (0.017)	0.140 (0.020)	0.012 (0.008)	[-0.003,0.028]
1982	0.170 (0.017)	0.145 (0.018)	0.141 (0.024)	0.011 (0.009)	[-0.008,0.028]
1983	0.192 (0.017)	0.154 (0.018)	0.160 (0.021)	0.024 (0.009)	*[0.008,0.042]
1984	0.204 (0.016)	0.150 (0.019)	0.146 (0.023)	0.023 (0.010)	*[0.003,0.042]
1985	0.212 (0.017)	0.177 (0.019)	0.179 (0.021)	0.021 (0.009)	*[0.003,0.040]
1986	0.193 (0.020)	0.181 (0.020)	0.188 (0.021)	0.026 (0.010)	*[0.006,0.045]
1987	0.203 (0.018)	0.149 (0.019)	0.132 (0.020)	0.053 (0.010)	*[0.034,0.072]
1988	0.195 (0.015)	0.165 (0.018)	0.165 (0.023)	0.024 (0.009)	*[0.006,0.041]
1989	0.196 (0.018)	0.159 (0.019)	0.161 (0.021)	0.021 (0.010)	*[0.001,0.039]
1990	0.194 (0.014)	0.156 (0.017)	0.161 (0.019)	0.025 (0.009)	*[0.010,0.042]
1991	0.182 (0.016)	0.152 (0.017)	0.161 (0.021)	0.018 (0.010)	[-0.001,0.037]
1992	0.195 (0.015)	0.173 (0.016)	0.165 (0.019)	0.024 (0.009)	*[0.006,0.043]
1993	0.214 (0.017)	0.207 (0.017)	0.201 (0.019)	0.011 (0.009)	[-0.006,0.029]
MEAN	0.192	0.162	0.161	0.020	
p-value		0.000*	0.000*	0.000*	

– C.I. denotes the confidence interval for $\text{Cov}(C, \beta(X,V))/C_{mean}$

– Last line of the table contains p-values for the hypotheses: (on average over 20 years) $\beta_{agg} - \beta_{mean} = 0$, $\beta_{agg} - \beta = 0$, and $\text{Cov}(C, \beta(X,V))/C_{mean} = 0$, respectively.

– Asterisks denote rejection of equality of the aggregate elasticity and the average individual elasticity at the 5% level.

Table 3.1: Income elasticities of demand for 'food expenditure.'

year	$\hat{\beta}_{agg}$	$\hat{\beta}_{mean}$	$\bar{\beta}$	$\text{Cov}(\widehat{C, \beta(X}, V))/\bar{C}$	C.I.
1974	0.187 (0.043)	0.164 (0.031)	0.165 (0.036)	0.016 (0.017)	[-0.017,0.049]
1975	0.193 (0.048)	0.165 (0.033)	0.177 (0.033)	-0.010 (0.019)	[-0.046,0.025]
1976	0.128 (0.035)	0.158 (0.026)	0.158 (0.032)	-0.003 (0.015)	[-0.034,0.027]
1977	0.087 (0.033)	0.131 (0.028)	0.145 (0.030)	-0.022 (0.015)	[-0.052,0.011]
1978	0.195 (0.035)	0.195 (0.034)	0.205 (0.039)	0.002 (0.017)	[-0.035,0.035]
1979	0.173 (0.032)	0.167 (0.028)	0.163 (0.033)	-0.016 (0.014)	[-0.047,0.009]
1980	0.131 (0.041)	0.158 (0.030)	0.167 (0.031)	-0.014 (0.014)	[-0.042,0.010]
1981	0.097 (0.030)	0.109 (0.023)	0.114 (0.029)	-0.001 (0.013)	[-0.024,0.027]
1982	0.196 (0.032)	0.180 (0.027)	0.174 (0.030)	0.015 (0.013)	[-0.011,0.037]
1983	0.189 (0.027)	0.188 (0.025)	0.195 (0.027)	0.001 (0.012)	[-0.021,0.026]
1984	0.241 (0.030)	0.241 (0.023)	0.239 (0.028)	0.000 (0.012)	[-0.023,0.024]
1985	0.226 (0.028)	0.204 (0.024)	0.194 (0.027)	0.019 (0.011)	[-0.004,0.041]
1986	0.194 (0.031)	0.144 (0.025)	0.138 (0.028)	0.009 (0.012)	[-0.017,0.032]
1987	0.196 (0.023)	0.179 (0.022)	0.184 (0.022)	-0.004 (0.010)	[-0.024,0.018]
1988	0.174 (0.025)	0.163 (0.021)	0.172 (0.022)	0.005 (0.010)	[-0.015,0.024]
1989	0.176 (0.022)	0.170 (0.019)	0.173 (0.021)	-0.010 (0.009)	[-0.027,0.010]
1990	0.153 (0.024)	0.150 (0.019)	0.156 (0.022)	-0.001 (0.010)	[-0.021,0.019]
1991	0.131 (0.020)	0.121 (0.017)	0.118 (0.022)	0.002 (0.009)	[-0.016,0.019]
1992	0.154 (0.023)	0.142 (0.023)	0.154 (0.021)	0.001 (0.009)	[-0.015,0.017]
1993	0.176 (0.020)	0.158 (0.020)	0.154 (0.021)	0.004 (0.009)	[-0.013,0.019]
MEAN	0.170	0.164	0.167	0.000	
p-value		0.117	0.478	0.941	

- C.I. denotes the confidence interval for $\text{Cov}(C, \beta(X, V))/C_{mean}$
- Last line of the table contains p-values for the hypotheses: (on average over 20 years) $\beta_{agg} - \beta_{mean} = 0$, $\beta_{agg} - \beta = 0$, and $\text{Cov}(C, \beta(X, V))/C_{mean} = 0$, respectively.
- Asterisks denote rejection of equality of the aggregate elasticity and the average individual elasticity at the 5% level.

Table 3.2: Income elasticities of demand for 'fuel and light.'

year	$\hat{\beta}_{agg}$	$\hat{\beta}_{mean}$	$\bar{\beta}$	$\widehat{\mathrm{Cov}(C, \beta(X,V))}/\bar{C}$	C.I.
1974	0.953 (0.059)	0.845 (0.038)	0.814 (0.046)	0.077 (0.034)	*[0.015,0.143]
1975	1.042 (0.070)	0.883 (0.039)	0.855 (0.047)	0.084 (0.040)	*[0.004,0.167]
1976	0.936 (0.058)	0.811 (0.043)	0.804 (0.043)	0.025 (0.030)	[-0.036,0.089]
1977	0.903 (0.063)	0.806 (0.039)	0.799 (0.045)	0.070 (0.039)	[-0.005,0.147]
1978	0.900 (0.057)	0.805 (0.039)	0.780 (0.052)	0.061 (0.038)	[-0.009,0.129]
1979	1.129 (0.087)	0.823 (0.040)	0.775 (0.048)	0.084 (0.041)	*[0.008,0.171]
1980	0.875 (0.063)	0.653 (0.040)	0.622 (0.049)	0.133 (0.033)	*[0.068,0.207]
1981	0.989 (0.058)	0.773 (0.033)	0.745 (0.043)	0.056 (0.033)	[-0.011,0.122]
1982	0.872 (0.085)	0.715 (0.037)	0.697 (0.047)	0.119 (0.034)	*[0.052,0.184]
1983	1.018 (0.056)	0.771 (0.037)	0.735 (0.043)	0.122 (0.033)	*[0.051,0.184]
1984	0.935 (0.056)	0.759 (0.041)	0.750 (0.048)	0.151 (0.036)	*[0.079,0.223]
1985	1.052 (0.055)	0.734 (0.037)	0.706 (0.044)	0.195 (0.032)	*[0.138,0.262]
1986	1.031 (0.081)	0.689 (0.040)	0.671 (0.047)	0.176 (0.034)	*[0.108,0.240]
1987	0.894 (0.051)	0.707 (0.033)	0.692 (0.047)	0.262 (0.034)	*[0.195,0.327]
1988	0.990 (0.057)	0.747 (0.035)	0.731 (0.037)	0.264 (0.040)	*[0.192,0.342]
1989	0.964 (0.072)	0.708 (0.030)	0.690 (0.039)	0.144 (0.035)	*[0.084,0.207]
1990	0.939 (0.046)	0.758 (0.033)	0.746 (0.041)	0.124 (0.031)	*[0.057,0.189]
1991	1.103 (0.069)	0.773 (0.033)	0.754 (0.041)	0.181 (0.038)	*[0.111,0.260]
1992	1.079 (0.062)	0.679 (0.030)	0.653 (0.041)	0.117 (0.038)	*[0.047,0.196]
1993	0.988 (0.069)	0.725 (0.034)	0.716 (0.037)	0.163 (0.039)	*[0.083,0.243]
MEAN	0.980	0.758	0.737	0.130	
p-value		0.000*	0.000*	0.000*	

- C.I. denotes the confidence interval for $\mathrm{Cov}(C, \beta(X,V))/C_{mean}$
- Last line of the table contains p-values for the hypotheses: (on average over 20 years) $\beta_{agg} - \beta_{mean} = 0$, $\beta_{agg} - \beta = 0$, and $\mathrm{Cov}(C, \beta(X,V))/C_{mean} = 0$, respectively.
- Asterisks denote rejection of equality of the aggregate elasticity and the average individual elasticity at the 5% level.

Table 3.3: Income elasticities of demand for 'services.'

year	$\hat{\beta}_{agg}$	$\hat{\beta}_{mean}$	$\bar{\beta}$	$\widehat{\text{Cov}(C, \beta(X,V))}/\bar{C}$	C.I.
1974	0.805 (0.065)	0.828 (0.073)	0.810 (0.070)	-0.003 (0.052)	[-0.104,0.094]
1975	0.845 (0.069)	0.777 (0.074)	0.766 (0.072)	0.059 (0.062)	[-0.062,0.171]
1976	0.912 (0.099)	0.826 (0.072)	0.821 (0.081)	-0.084 (0.064)	[-0.204,0.037]
1977	0.822 (0.068)	0.902 (0.074)	0.883 (0.078)	-0.085 (0.061)	[-0.204,0.041]
1978	0.814 (0.065)	0.893 (0.070)	0.864 (0.075)	-0.033 (0.058)	[-0.138,0.082]
1979	0.714 (0.070)	0.716 (0.076)	0.725 (0.079)	-0.003 (0.062)	[-0.120,0.112]
1980	0.690 (0.061)	0.776 (0.070)	0.753 (0.073)	-0.061 (0.050)	[-0.157,0.044]
1981	0.646 (0.060)	0.613 (0.067)	0.607 (0.081)	0.022 (0.053)	[-0.085,0.127]
1982	0.757 (0.060)	0.749 (0.067)	0.732 (0.074)	0.023 (0.053)	[-0.077,0.126]
1983	0.688 (0.060)	0.666 (0.071)	0.651 (0.072)	0.049 (0.056)	[-0.064,0.162]
1984	0.786 (0.066)	0.733 (0.070)	0.733 (0.076)	0.008 (0.053)	[-0.094,0.111]
1985	0.760 (0.060)	0.696 (0.068)	0.701 (0.073)	0.034 (0.057)	[-0.067,0.152]
1986	0.785 (0.068)	0.721 (0.066)	0.720 (0.068)	0.008 (0.052)	[-0.090,0.110]
1987	0.716 (0.051)	0.691 (0.058)	0.691 (0.060)	0.030 (0.045)	[-0.058,0.118]
1988	0.700 (0.053)	0.669 (0.057)	0.670 (0.064)	-0.027 (0.046)	[-0.114,0.058]
1989	0.720 (0.054)	0.670 (0.059)	0.679 (0.064)	-0.009 (0.049)	[-0.109,0.088]
1990	0.651 (0.054)	0.538 (0.055)	0.545 (0.064)	0.008 (0.048)	[-0.097,0.100]
1991	0.640 (0.052)	0.696 (0.056)	0.707 (0.065)	-0.048 (0.050)	[-0.146,0.052]
1992	0.594 (0.052)	0.593 (0.054)	0.597 (0.064)	0.008 (0.048)	[-0.084,0.103]
1993	0.737 (0.057)	0.638 (0.056)	0.648 (0.059)	0.066 (0.048)	[-0.029,0.159]
MEAN	0.739	0.720	0.715	-0.002	
p-value		0.145	0.086	0.911	

- C.I. denotes the confidence interval for $\text{Cov}(C, \beta(X,V))/C_{mean}$
- Last line of the table contains p-values for the hypotheses: (on average over 20 years) $\beta_{agg} - \beta_{mean} = 0$, $\beta_{agg} - \beta = 0$, and $\text{Cov}(C, \beta(X,V))/C_{mean} = 0$, respectively.
- Asterisks denote rejection of equality of the aggregate elasticity and the average individual elasticity at the 5% level.

Table 3.4: Income elasticities of demand for 'clothing and footware.'

year	$\hat{\beta}_{agg}$	$\hat{\beta}_{mean}$	$\bar{\beta}$	$\widehat{\text{Cov}(C, \beta(X, V))}/\bar{C}$	C.I.
1974	0.538 (0.023)	0.476 (0.019)	0.474 (0.023)	0.038 (0.011)	*[0.014,0.058]
1975	0.551 (0.021)	0.486 (0.017)	0.482 (0.022)	0.041 (0.012)	*[0.019,0.064]
1976	0.528 (0.025)	0.458 (0.019)	0.453 (0.023)	0.072 (0.011)	*[0.049,0.092]
1977	0.488 (0.022)	0.432 (0.019)	0.441 (0.021)	0.051 (0.011)	*[0.028,0.073]
1978	0.540 (0.023)	0.497 (0.019)	0.497 (0.024)	0.027 (0.013)	*[0.002,0.053]
1979	0.526 (0.023)	0.465 (0.019)	0.464 (0.023)	0.033 (0.012)	*[0.011,0.058]
1980	0.510 (0.021)	0.435 (0.018)	0.434 (0.023)	0.059 (0.011)	*[0.038,0.081]
1981	0.517 (0.021)	0.423 (0.018)	0.423 (0.019)	0.044 (0.011)	*[0.023,0.068]
1982	0.509 (0.028)	0.432 (0.018)	0.434 (0.021)	0.054 (0.011)	*[0.033,0.075]
1983	0.509 (0.023)	0.432 (0.019)	0.435 (0.024)	0.062 (0.011)	*[0.038,0.086]
1984	0.546 (0.026)	0.472 (0.019)	0.470 (0.025)	0.054 (0.011)	*[0.031,0.075]
1985	0.564 (0.020)	0.467 (0.018)	0.456 (0.021)	0.079 (0.011)	*[0.057,0.101]
1986	0.568 (0.025)	0.451 (0.018)	0.438 (0.024)	0.071 (0.014)	*[0.046,0.101]
1987	0.550 (0.019)	0.443 (0.016)	0.441 (0.019)	0.072 (0.012)	*[0.051,0.094]
1988	0.573 (0.020)	0.477 (0.016)	0.465 (0.021)	0.063 (0.010)	*[0.046,0.082]
1989	0.517 (0.022)	0.452 (0.017)	0.454 (0.021)	0.050 (0.011)	*[0.029,0.073]
1990	0.520 (0.017)	0.436 (0.015)	0.435 (0.019)	0.055 (0.010)	*[0.035,0.075]
1991	0.540 (0.018)	0.463 (0.015)	0.453 (0.020)	0.050 (0.011)	*[0.028,0.072]
1992	0.509 (0.034)	0.449 (0.015)	0.449 (0.020)	0.038 (0.016)	*[0.013,0.076]
1993	0.568 (0.024)	0.455 (0.018)	0.459 (0.020)	0.086 (0.013)	*[0.061,0.109]
MEAN	0.534	0.455	0.453	0.055	
p-value		0.000*	0.000*	0.000*	

- C.I. denotes the confidence interval for $\text{Cov}(C, \beta(X, V))/C_{mean}$
- Last line of the table contains p-values for the hypotheses: (on average over 20 years) $\beta_{agg} - \beta_{mean} = 0$, $\beta_{agg} - \beta = 0$, and $\text{Cov}(C, \beta(X, V))/C_{mean} = 0$, respectively.
- Asterisks denote rejection of equality of the aggregate elasticity and the average individual elasticity at the 5% level.

Table 3.5: Income elasticities of demand for 'total (nondurable) expenditure.'

provides further evidence for the appropriateness of our crucial assumption and our estimation strategy.

5) The perhaps most interesting empirical result is that aggregate elasticity can be very different from the mean of individual elasticities. The magnitude of this difference varies from commodity to commodity. For expenditure on food and services, as well as for total expenditure, aggregate income elasticity is greater than the mean individual elasticity for all sample years. In the extreme case of expenditure on services, the difference can be as large as 30% of the aggregate elasticity. On the other hand, for the commodity groups 'clothing and footware' and 'fuel and light' aggregate and mean individual elasticities are quite close.

As mentioned in the last section, in order to assess whether the discrepancy between β_{agg} and β_{mean} is statistically significant we expose this difference to several tests. The p-values from the Wilcoxon test for matched pairs of the hypothesis that the average (over the period 1974-1993) difference between β_{agg} and β_{mean} is zero are given in the last line of the table. The last line of the first and the second column report p-values based on the comparison of $\hat{\beta}_{agg}$ with $\hat{\beta}_{mean}$ and $\bar{\beta}$, respectively. Asterisks in Tables 3.1-3.5 denote the significance at the 95% level. According to these p-values, the aggregate elasticity is significantly greater than the mean of individual elasticities for 'food', 'services', and total expenditure. For 'clothing and footware' and 'fuel and light' the difference between β_{agg} and β_{mean} is not significant.

Similarly, according to the Wilcoxon signed-rank test of the hypothesis $\mathrm{Cov}(C, \beta(X, V))/C_{mean} = 0$, we reject it in favor of $\mathrm{Cov}(C, \beta(X, V))/C_{mean} > 0$ for 'food', 'services', and total expenditure. For 'fuel and light' and 'clothing and footware' the covariance term is not significantly different from zero.

The discussion above regards the average difference between the aggregate elasticity and the mean individual elasticity over the sample period of 20 years. We perform a bootstrap to assess the statistical significance of this difference for each year of the sample. The figures in the last column of the Tables 3.1-3.5 are the 95% bootstrap confidence interval for $\mathrm{Cov}(\widehat{C, \beta(X}, V))/\bar{C}$, which we can use to test

$H_0 : \mathrm{Cov}(C, \beta(X, V))/C_{mean} = 0$. The main result of this test is that for expenditure on 'food', 'services', and 'total expenditure' the covariance term is significantly positive for almost all sample years. For the remaining commodity groups 'clothing and footware' and 'fuel and light' the covariance term is not significant.

year	food	fuel	services	clothing
1974	0.328 (0.038)	0.346 (0.062)	1.863 (0.095)	1.742 (0.111)
1975	0.379 (0.042)	0.331 (0.061)	1.9431 (0.098)	1.656 (0.099)
1976	0.384 (0.039)	0.340 (0.072)	2.0446 (0.110)	1.979 (0.118)
1977	0.304 (0.032)	0.430 (0.070)	2.0289 (0.126)	2.052 (0.125)
1978	0.393 (0.045)	0.390 (0.057)	1.922 (0.127)	1.812 (0.126)
1979	0.287 (0.046)	0.435 (0.056)	1.7942 (0.110)	1.784 (0.098)
1980	0.366 (0.040)	0.369 (0.053)	1.681 (0.122)	1.827 (0.123)
1981	0.359 (0.053)	0.291 (0.052)	2.1376 (0.112)	1.686 (0.140)
1982	0.279 (0.047)	0.390 (0.063)	1.7013 (0.086)	1.773 (0.143)
1983	0.329 (0.049)	0.404 (0.066)	2.0172 (0.114)	1.594 (0.139)
1984	0.313 (0.042)	0.490 (0.063)	1.7483 (0.113)	1.818 (0.145)
1985	0.357 (0.048)	0.382 (0.045)	1.6626 (0.115)	1.645 (0.123)
1986	0.463 (0.048)	0.376 (0.075)	1.9541 (0.086)	1.957 (0.135)
1987	0.312 (0.061)	0.487 (0.073)	1.6134 (0.130)	1.593 (0.112)
1988	0.261 (0.042)	0.346 (0.046)	1.796 (0.109)	1.726 (0.121)
1989	0.390 (0.045)	0.363 (0.056)	1.556 (0.104)	1.659 (0.103)
1990	0.382 (0.048)	0.329 (0.050)	1.8716 (0.095)	1.452 (0.096)
1991	0.388 (0.030)	0.260 (0.046)	1.7051 (0.086)	1.744 (0.093)
1992	0.387 (0.044)	0.311 (0.051)	1.6847 (0.081)	1.514 (0.104)
1993	0.308 (0.041)	0.283 (0.048)	1.6436 (0.092)	1.613 (0.094)
MEAN	0.348	0.368	1.818	1.731

Table 3.6: Estimates of budget elasticities $\beta_{mean,tot}$ for all commodity groups.

It is important to note that estimation results in Tables 3.1-3.5 are elasticities with respect to income. However, under additional assumptions described in Section 3.2 our methodology allows estimation of elasticities with respect to budget. For the sake of completeness and comparability with other studies we present estimates of the mean of individual budget elasticities $\beta_{mean,tot}$ for several commodity groups in Table 3.6. It is not surprising that these estimates are substantially greater than the corresponding mean of income elasticities as the latter do not take the savings behavior into account. Indeed, $\hat{\beta}_{mean,tot}$ is roughly twice as large as the corresponding

$\hat{\beta}_{mean}$, which seems intuitive as $\hat{\beta}_{mean}$ for total expenditure is approximately equal to 0.5.

To sum up, we found strong empirical evidence for aggregate elasticity to be greater than the mean of individual elasticities for commodity groups 'food', 'services' and total expenditure. In contrast, for commodity groups 'fuel and light' and 'clothing and services' the aggregate elasticity seems neither to overestimate, nor to underestimate the average individual elasticity.

The above result has extensive implications for both policy makers and applied researchers. As for the former, the knowledge of the relationship between the aggregate elasticity and the distribution of individual elasticities is crucial for correct evaluation of economic reforms. For instance, if one wants to assess possible changes in demand due to an income tax reform, one should take heterogeneity in income elasticities into account. For the latter, it is important to know that one should not interpret the aggregate elasticity in terms of mean individual elasticities, since the difference between them can be of magnitude of even 30% of the aggregate elasticity.

References

Abramovitz, M. (1956): "Resource and Output Trends in the United States since 1870", *American Economic Review* 46, 5-23.

Acemoglu, D. and F. Zilibotti (2001): "Productivity Differences", *Quarterly Journal of Economics* 116, 563-606.

Aghion, P. and R. Griffith (2005): "Competition and Growth", *MIT Press* , Cambridge, MA.

Aitchison, J. and C.G.G. Aitken (1976): "Multivariate Binary Discrimination by Kernel Method", *Biometrika* 63, 413–420.

Antonelli, G.B. (1886): "Sulla teoria matematica della economia politica", Pisa: Nella Tipografia del Folchetto (translated by J.S. Chipman and A.P. Kirman in "Preferences, Utility and Demand", ed. J.S. Chipman et al. New York: Harcourt Brace Jovanovich, 333-60.)

Arns, J. (2006): "Evolutions of Age and Income Distributions in the UK: Empirical Analysis of Data Consistency and Functional Modelling," PhD Dissertation, Universität Mainz.

Arns, J. and K. Bhattacharya (2005): "Modelling Aggregate Consumption Growth with Time-Varying Parameters", *Bonn Econ Discussion Papers* 15/2005, Universität Bonn.

Atkinson, A.B. (1970): "On the Measurement of Inequality", *Journal of Economic Theory* 2, 244-263.

Banks, J., Blundell, R.W., and A. Lewbel (1997): "Quadratic Engel Curves, Indirect Tax Reform, and Welfare Measurement", *Review of Economics and Statistics* 79, 527-539.

Basu, S. and D. Weil (1998): "Appropriate Technology and Growth", *Quarterly Journal of Economics* 113, 1025-1054.

Blundell, R.W., Pashardes, P., and G. Weber (1993): "What do we Learn About Consumer Demand Patterns from Micro Data?", *American Economic Review* 83, 570-597.

Caselli, F. (2005): "Accounting for Cross-Country Income Differences", *Handbook of Economic Growth*, Edition 1, Volume 1, Chapter 9, Elsevier, 679-741.

Campbell, J.Y. and G. Mankiw (1990): "Permanent Income, Current Income, and Consumption", *Journal of Business and Economic Statistics* 8, 265-279.

Chakrabarty, M., Schmalenbach, A., and J. S. Racine (2006): "On the Distributional Effects of Income in an Aggregate Consumption Relation", *Canadian Journal of Economics* 39, 1221–1243.

Chow, G. (1983): "Econometrics", Chapter 9, McGraw-Hill Book Company, New York.

Cowell, F.A, Jenkins, S.P., and J.A. Litchfield (1996): "The Changing Shape of the UK Income Distribution: Kernel Density Estimates", in J.Hills (ed) "New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom", Cambridge University Press, Cambridge.

Davidson, J.E.H., Hendry, D.F., Srba, F., and S. Yeo (1978): "Econometric Modelling of the Aggregate Time-Series Relationship between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal* 88, 661-692.

Deaton, A.S. and J. Muellbauer (1980): *Economics and Consumer Behavior*. Cambridge University Press.

Fan, I. and I. Gijbels (1996): "Local Polynomial Modeling and Its Applications", Chapman and Hall, London.

Felipe, J. and F.M. Fisher (2003): "Aggregation in Production Functions: What Applied Economists Should Know", *Metroeconomica* 54, 208-262.

Fisher, F.M. (1993): "Aggregation. Aggregate Production Functions and Related Topics", MIT Press, Cambridge, MA.

Ginn, J. and Arber, S. (2000): "Ethnic Inequality in Later Life: Variation in Financial Circumstances by Gender and Ethnic Group", *Education and Ageing* 15, 65-83.

Gorman, W.M. (1953): "Community Preference Fields", *Econometrica* 21, 63-80.

Gottschalk, P. and T.M. Smeeding (2001): "Empirical Evidence on Income Inequality in Industrial Countries", in "Handbook of Income Distribution", Atkinson, A.B., Bourguignon, F. (eds.), Amsterdam, Elsevier, 261-304.

Hansen, E.B. (1992): "Testing for Parameter Instability in Linear Regression Models", *Journal of Policy Modeling* 14, 517-533.

Hall, P., Racine, J.S., and Q. Li (2004): "Cross-Validation and the Estimation of Conditional Probability Densities", *Journal of the American Statistical Association* 99, 1015–1026.

Hall, R. and C. Jones (1999): "Why Do Some Countries Produce So Much More Output than Others?", *Quarterly Journal of Economics* 114, 83-116.

Härdle, W. and J.D. Hart (1992): "A Bootstrap Test for a Positive Definiteness of Income Effect Matrices", *Econometric Theory* 8, 276–290.

Härdle, W. and T.M. Stoker (1989): "Investigating Smooth Multiple Regression by the Method of Average Derivatives", *Journal of the American Statistical Association* 84, 986–995.

Hildenbrand, W. and A. Kneip (1999): "Demand Aggregation under Structural Stability", *Journal of Mathematical Economics* 31, 81-110.

Hildenbrand, W., Kneip, A., and K. Utikal (1999): "Une analyse non paramétrique des distributions du revenu et des caractéristiques des ménages", *Revue de Statistique Appliquée* 47, 39-56.

Hildenbrand, W. and A. Kneip (2005): "Aggregate Behavior and Microdata", *Games and Economic Behavior* 50, 3–27.

Hoderlein, S. and E. Mammen (2007): "Identification of Marginal Effects in Nonseparable Models without Monotonicity", *Econometrica*, (to appear in 2007).

Hopenhayn, H. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibria", *Econometrica* 60, 1127-1150.

Houthakker, H.S. (1957): “An International Comparison of Household Expenditure Patterns, Commemorating the Century of Engel’s Law”, *Econometrica* 25, 532-551.

Jorgenson, D.W., Lau, L.J. and T.M. Stoker (1982): “The Transcendental Logarithmic Model of Aggregate Consumer Behavior”, in R. Basman and G. Rhodes, eds., *Advances in Econometrics*, Greenwich: JAI Press.

Jorgenson, D.W. (2005): “Accounting for Growth in the Information Age”, *Handbook of Economic Growth*, Edition 1, Volume 1, Chapter 10, Elsevier 743-815.

Keller, W. (2004): “International Technology Diffusion”, *Journal of Economic Literature* 152, 752-782.

Kemsley, W., Redpath, R., and M. Holmes (1980): *Family Expenditure Survey Handbook*, London: HMSO.

Koenker, R. (1994): “Confidence Intervals for Quantile Regression”, in Petr Mandl and Marie Huskova, eds., *Proceedings of the 5th Prague Symposium on Asymptotic Statistics*, Heidelberg: Physica-Verlag.

Levinsohn, J. and A. Petrin (2003): “Estimating Production Functions Using Inputs to Control for Unobservables”, *Review of Economic Studies* 70, 317-341.

Lewbel, A. (1989): “Nesting the Aids and Translog Demand Systems”, *International Economic Review* 30, 349–356.

Lewbel, A. (1990): “Income Distribution Movements and Aggregate Money Illusions”, *Journal of Econometrics* 43, 35–42.

Lewbel, A. (1992): “Aggregation with Log-linear Models”, *Review of Economic Studies* 59, 635–642.

Li, Q. (1996): “Nonparametric Testing of Closeness between Two Unknown Distribution Functions”, *Econometric Reviews*, 15, pp. 261-274.

Li, Q., Maasoumi, E., and J.S. Racine (2007): “A Nonparametric Test for Equality of Distributions with Mixed Categorical and Continuous Data”, *mimeo*.

Li, Q. and J.S. Racine (2004): “Cross-Validated Local Linear Nonparametric Regression”, *Statistica Sinica* 14, 485–512.

Li, Q. and J.S. Racine (2006): "Nonparametric Estimation of Conditional CDF and Quantile Functions with Mixed Categorical and Continuous Data", *Journal of Business and Economic Statistics.*

Malinvaud, E. (1993): "A Framework for Aggregation Theories", *Ricerche Economiche* 47, 107-135.

Mallows, C.L. (1973): "Some Comments on C_p", *Technometrics* 15, 661-675.

Matzkin, R.L. (2003): "Nonparametric Estimation of Nonadditive Random Functions", *Econometrica* 71, 1332–1375.

Melitz, M. (2003): "The Impact of Trade on intra-industry Reallocations and Aggregate Productivity", *Econometrica* 71, 1695-1725.

Nataf, A. (1948): "Sur la possibilité de construction de certains macromodèles", *Econometrica* 16, 232-244.

Nicoletti, G. and S. Scarpetta (2003): "Regulation, Productivity and Growth: OECD Evidence", *Economic Policy* 18, 9-72.

Olley, G. and A. Pakes (1996): "The Dynamics of Productivity in the Telecomunications Equipment Industry", *Econometrica* 64, 1263-1279.

Palis J. and S. Smale (1970): "Structural Stability Theorems", *Proceedings of the Institute on Global Analysis*, American Mathematical Society 14, 223-232.

Paluch, M. (2004): "A Note on the Structural Stability of the Joint Distribution of Income and Wealth", *Bonn Econ Discussion Paper* 16/2004, Universität Bonn.

Paluch, M., Kneip, A., and W. Hildenbrand (2007): "Individual versus Aggregate Income Elasticities for Heterogeneous Populations", *Bonn Econ Discussion Paper* 13/2007, Universität Bonn.

Paluch, M., and M. Schiffbauer (2007): "Distributional Effects of Capital and Labor on Economic Growth",*mimeo*, Universität Bonn.

Pareto, V. (1896-1897): "Cours d'économie politique", Vol. 1-2, Rouge, Lausanne.

Piachaud, D. and H. Sutherland (2002): "Changing Poverty post-1997", Centre for Analysis of Social Exclusion, London School of Economics, *CASEPaper* 63.

Pittau, M.G. and R. Zelli (2001): "Income Distribution in Italy: A Nonparametric Analysis", *Statistical Methods and Applications* 10, 175-189.

Portnoy, S. and R. Koenker (1997): "The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators", *Statistical Science* 12, 279–300.

Quah, D. (1997): "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth* 2, 27-59.

Prescott, E. (1998): "Needed: A Theory of Total Factor Productivity", *International Economic Review* 39, 525-551.

Racine, J.S. and Q. Li (2004): "Nonparametric Estimation of Regression Functions with both Categorical and Continuous Data", *Journal of Econometrics* 119, 99–130.

Ramsey, J.B. (1969): "Tests for Specification Error in Classical Linear Least Squares Regression Analysis", *Journal of the Royal Statistical Society, Series B,* 31, 350–371.

Romer, D. (2005): "Advanced Macroeconomics", Third Edition. McGraw-Hill/Irwin.

Sala-i-Martin, X. (2002): "The World Distribution of Income (Estimated from Individual Country Distributions)", *NBER Working Paper* 8933

Schumpeter, J. (1949): "Vilfredo Pareto, 1948-1923", Reprinted in J. Schumpeter (1951), "Ten Great Economists", Oxford University Press, Oxford.

Sheather, S.J. and M.C. Jones (1991): "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation", *Journal of the Royal Statistics Association, Series B* 53, 683-690.

Solow, R. (1956): "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics* 70, 65-94.

Stigler, G.J. (1954): "The Early History of Empirical Studies of Consumer Behavior", *The Journal of Political Economy* 62, 95–113.

Wang, M.-C. and J. van Ryzin (1981): "A Class of Smooth Estimators for Discrete Estimation," *Biometrika* 68, 301–309.

Wilcoxon, F. (1945): "Individual Comparisons by Ranking Methods", *Biometrics Bulletin* 1, 80–83.

Yu, K. and M.C. Jones (1998): "Local Linear Quantile Regression", *Journal of the American Statistical Association* 93, 228–237.