Labour Supply and the "Law of Demand"

Thesis submitted for the degree of Ph.D.

Thomas Philipp

London School of Economics and Political Science
University of London

January 1994

Abstract

The well-known "law of supply and demand" says that an increase in the price of a commodity leads to a decrease in the aggregate demand for this commodity and an increase in aggregate supply. There is, however, no theoretical foundation for this "law". Empirical evidence, on the other hand, should be interpreted with care. If one estimates the parameters of certain functional forms for demand and supply functions, then the results may simply be consequences of the parametric assumptions made in estimation.

The first chapter of the thesis discusses the implications of the assumption of profit and utility maximisation for the properties of demand and supply functions. It explains why economic rationality on the micro-level does not, in general, lead to macroeconomic regularities and suggests replacing the consumption sector of the neoclassical equilibrium model by a large population of individually small consumers.

Such a population will be explored in the second chapter. The chapter is a direct outgrowth of a basic contribution by W. Hildenbrand: "On the Law of Demand", Econometrica 1983. In W. Hildenbrand's model the market demand function is defined by integrating an individual demand function with respect to an exogenously given income distribution. We build into the model an individual labour supply function and then compare the matrix of aggregate income effects studied by W. Hildenbrand with that obtained by integrating the individual demand function with respect to a distribution of wage rates.

The empirical part of the thesis analyses the labour supply and earnings data in the U.K. Family Expenditure Survey 1970-85. Using non-parametric smoothing methods, the elasticity of labour supply with respect to the wage rate is estimated for several groups of workers. The estimations for full-time workers confirm the famous "downward sloping" labour supply function. The estimated elasticities for the entire population of workers for the years 1970-85 have the mean value 0.2 and the standard deviation 0.02.

Preface

I would like to thank my dissertation advicer Anthony B. Atkinson for drawing my attention to labour supply and for his suggestions. In particular, I would like to thank him for his patience. The thesis was written over a period which was longer than I wish to remember.

My debt of gratitude to the following two people goes beyond the thesis. Joanna Gomulka told me what I had to know in order to carry out the empirical part of the thesis. Rolf Peter Schneider proof-read the two empirical chapters. Of course, any errors are my own responsibility.

The data which will be explored in Chapters 2 and 3 were made available by the Social Science Research Council Survey Archive at the University of Essex. The estimations were carried out in the academic year 1989-90 at the Suntory-Toyota International Centre for Economics and Related Disciplines of the London School of Economics.

The thesis was written under the auspices of the European Doctoral Programme in Quantitative Economics. I would like to express my gratitude that this institution gave me the opportunity to study at the London School of Economics.

Finally, I gratefully acknowledge the financial support of the German Academic Exchange Service and the Volkswagen Foundation.

Düsseldorf, January 1994 Thomas Philipp

Table of Contents

Notation and Definitions

We will denote the set of real numbers by R; the set of all non-negative real numbers will be denoted by R_+, and the set of all positive real numbers will be denoted by R_{++}. The sets R^n, R^n_+ and R^n_{++} are defined as the n-fold Cartesian products of R, R_+ and R_{++}, respectively.

The sum of two vectors $x=(x_1,\ldots,x_n)$ and $y=(y_1,\ldots,y_n)$ is defined as $x+y = (x_1+y_1,\ldots,x_n+y_n)$, and the product of x by a real number α is defined as $\alpha x = (\alpha x_1,\ldots,\alpha x_n)$.

Let x, y, $x^1,\ldots,x^m$ be elements of R^n. For the sum $x^1+\ldots+x^m$ we write Σx^i. The *scalar product* of x and y is denoted by xy and is given by $xy = \Sigma x_i y_i$, where $x=(x_1,\ldots,x_n)$ and $y=(y_1,\ldots,y_n)$; $||x||$ denotes the *Euclidean norm* of x, i.e., $||x||=\sqrt{xx}$. The symbol #A stands for the number of elements in a finite set A.

f: $D \longrightarrow R^m$ denotes a function f with domain D in R^n and range in R^m; f(A) is the set of all points f(x) such that $x \in A \subseteq D$. The components of f are indicated by $f_1,\ldots,f_m$. We denote the *Jacobian matrix* of f at $x=(x_1,\ldots,x_n)$ by $\delta f(x)$; the elements of $\delta f(x)$ will be denoted by $\delta_{x_j} f_i(x)$, i.e., $\delta_{x_j} f_i(x)$ is the partial derivative of f_i with respect to the j-th variable at x. If we want to emphasise that the derivative is evaluated at some particular point x^*, we write $\delta_{x_j} f_i(x)_{|x=x^*}$. The integral of f with respect to a measure μ on D is defined as

$$\int f d\mu = \left(\int f_1 d\mu,\ldots,\int f_m d\mu\right).$$

The function f: $D \longrightarrow R^m$ is called *homogeneous of degree zero* if $f(x) = f(\alpha x)$ for all $x \in D$ and all $\alpha > 0$ such that $\alpha x \in D$. A function g: $R_+ \longrightarrow R_+$ is said to be a *density* if g integrates to one, i.e., $\int g(x)dx = 1$.

A real $n \times n$ matrix $A=(a_{ij})$ is said to be *positive* (resp. *negative*) *semi-definite* if $xAx \geq 0$ (resp. $xAx \leq 0$) for all x in R^n, where $xAx=\Sigma a_{ij} x_i x_j$ (note that we do not assume that A is symmetric).

Chapter 0

Market Excess Demand Functions in General Equilibrium Theory

1. Topic of the Thesis

If one wants to study how a market economy responds to changes in its exogenous parameters, certain properties of the commodity demand and the labour supply function are required. Textbooks on macroeconomics usually assume that aggregate labour supply is an increasing function of the real wage; in virtually all partial equilibrium studies it is assumed that the demand for an aggregated commodity is a decreasing function of the commodity price. There is, however, no microeconomic foundation for assumptions of this type. Typically there will be individuals in the economy who respond to a price increase by increasing their demand. If such people are in the majority, the market demand function will not be monotone decreasing in the commodity price. The question arises whether it is possible to identify "broad" classes of distributions of consumption characteristics which lead to macroeconomic regularities. More precisely, are there testable (and not too restrictive) hypotheses on the distribution of personal characteristics which imply that the aggregate commodity demand (resp. labour supply) function has specific properties? An important step to an answer of this question was taken by W. Hildenbrand (1983).

Empirical evidence on the dependence of commodity demand and labour supply upon prices and wages should be interpreted with care. It is standard practice in the literature to estimate the parameters of certain functional forms for aggregate commodity demand and labour supply relationships. Observed regularities may therefore simply be consequences of the parametric assumptions made in estimation. For the case of the commodity demand function the problem of estimation was recently addressed by K. Hildenbrand and W. Hildenbrand (1986) and W. Hildenbrand (1989a).

The present thesis builds on the above three contributions. Chapter 1 emphasises the importance of the labour market for structural properties of the market demand function; the chapter discusses Hildenbrand (1983) and extends the model. Chapters 2 and 3 present an analysis of the earnings and labour supply data in the U.K. Family Expenditure Survey 1970-85. Chapter 2 is concerned with the distribution of wages and hours of work. In Chapter 3 labour supply curves for several populations of workers will be estimated. The novelty of our empirical study is that we use nonparametric smoothing techniques in order to estimate the elasticity of labour supply with respect to the wage rate. The chapters are written in such a way that they can be read independently of each other.

This chapter discusses the neoclassical equilibrium model. Section 2 gives an informal description of the model. Section 3 reviews the implications of the paradigm of profit and preference maximisation for the properties of demand and supply functions; we close the section with the results on excess demand functions by Debreu (1974) and Mantel (1976). In Section 4 we sketch the proof of Debreu's indeterminacy theorem.

2. The Neoclassical Equilibrium Model

The primitive concepts of the model are commodities, prices, technologies and preferences. There are two types of economic agents: consumers and producers. Consumers are characterised by preferences and income; producers are characterised by technologies. A certain behaviour of consumers and producers is assumed. Finally, an equilibrium concept is introduced. The model was first formulated by Walras (1874); the rigorous mathematical foundation was provided by Arrow and Debreu (1954). Excellent textbooks are, e.g., (in increasing order of abstraction) Varian (1984), Malinvaud (1972), Arrow and Hahn (1971) and Debreu (1959).

It is assumed that there is a finite number n of commodities. Commodities are labelled in such a manner that one can speak of commodity 1, commodity 2 and so on; the same applies to comsumers and producers. A *commodity bundle* is a collection $x=(x_1,\ldots,x_n)$ of the n goods, where x_i is to be read as "$|x_i|$ units of commodity i" (we will see below that some of the elements of x typically have a negative sign). To each good i a price $p_i>0$ is assigned; a list $p=(p_1,\dots,p_n)$ of the n prices is called a *price system*. The price p_i is interpreted as the amount which has to be paid today by an agent for one unit of commodity i which will be made available to him (in the future). Some of the n commodities are usually specific types of labour; hence some of the prices p_i are usually wage rates. (Notice that the model does not explain who determines the price system.)

A producer (resp. firm) uses inputs in order to produce certain outputs. A collection $y=(y_1,\dots,y_n)$ of amounts of inputs and outputs is called a *feasible production plan* if the outputs can be produced with the inputs. To be able to distinguish between inputs and outputs in a production plan, inputs have a negative sign and outputs have a positive sign. Thus, if $y_i<0$ then $-y_i$ units of commodity i are used as an input; if $y_i>0$ then y_i units of commodity i are produced; if $y_i=0$ then commodity i is not used in the production process described by y. The set of all feasible production plans, called the producer's *technology*, is a subset of R^n and will be denoted by Y.

Suppose the firm realises the production plan y in Y. Then the scalar product py represents its profit with respect to the price system p. It is assumed that the firm wants to maximise py, i.e., the production target is to choose a point y* in Y such that $py \le py^*$ for all y in Y, where p is exogenously given. Under this assumption, a firm is completely described by its technology. (Notice that we do not really describe a firm but merely technological knowledge. The classic article on the "theory of the firm" is Coase, 1937; for a survey see, e.g., Holmström and Tirole, 1989.)

The role of the consumer (resp. household) is to supply labour and to consume commodities, i.e., the consumer chooses a *consumption plan* x = $(x_1,\dots,x_n)$ in his *consumption set* X; X is the set of all consumption and labour supply combinations which the consumer could realise in principle (i.e., X is the consumer's "technology"). We reverse the above sign convention: the labour supply of a consumer has a negative sign, and his consumption is represented by positive numbers. From the point of view of the model, the crucial difference between a consumer and a producer is that the consumer does not want to maximise his income but his satisfaction which he derives from consumption; furthermore the firms are owned by the consumers (see Koopmans', 1957, first essay for a beautiful discussion of Robinson Crusoe's complex decision problems).

Given any two consumption plans x and x' in X, it is assumed that the consumer is able to say whether he likes x more than x', or vice versa, or whether he is indifferent between x and x'; furthermore, if from the point of view of the consumer x is at least as good as x' and x' is at least as good as x", then it is assumed that the consumer does not desire x" more than x. Hence, the tastes of a consumer can be described by a binary relation $\precsim$ on X, called *preference relation*, which is complete and transitive (or, equivalently, by a *utility function* u: $X \longrightarrow R$). The expression $x \precsim x'$ [resp. the inequality $u(x) \leq u(x')$] is interpreted as "the consumer does not like x more than x'"; if $x \precsim x'$ and $x' \precsim x$, we say "x is indifferent to x'". It is assumed that $\precsim$ does not depend on prices and on other consumers' tastes.

Suppose the consumer has non-labour income m. Then he can only realise consumption plans x which have the property that px does not exceed m. It is assumed that the consumer chooses a point x* in X such that $px^* \leq m$ and $x \precsim x^*$ for all x in X satisfying $px \leq m$ (that is to say, the consumer is a "fully rational" person).

Under certain assumptions on preferences (resp. utility functions) one can show the following: (i) x* exists and is uniquely determined by p and m; (ii) the function $x^*=f(p,m)$ is continuous (resp. differentiable); and (iii) the consumer does not keep any money back, i.e., $pf(p,m)=m$. Under certain assumptions on technologies one can show that a uniquely determined profit maximising production plan $y^*=y(p)$ exists and that y(p) is a continuous (resp. differentiable) function.

The function f(p,m) is called the *demand function* of the consumer; the function y(p) is called the *supply function* of the producer. (Note that f(p,m) contains the labour supply of the consumer at (p,m); y(p) contains the demand for labour of the firm at the price system p.)

There are H consumers in the economy having consumption sets $X_1,\ldots,X_H$ and preference relations $\precsim_1,\ldots,\precsim_H$; the demand of consumer i at (p,m) will be denoted by $f^i(p,m)$ $(i=1,\ldots,H)$. There are F firms having technologies

$Y_1,\dots,Y_F$; the supply of producer j at the price system p will be denoted by $y^j(p)$ $(j=1,\dots,F)$.

It remains to explain where non-labour income m stems from. In a private ownership economy all natural resources and all firms are owned by the consumers; furthermore, the consumers own various commodities which have been produced in the past. The *total resources* of the economy are represented by a commodity bundle $w=(w_1,\dots,w_n)$, i.e., w includes the natural resources of the society and commodities which have been produced in the past and which are still available.

The access to total resources which consumer i has is represented by a commodity bundle $w^i=(w^i_1,\dots,w^i_n)$; w^i is called the *initial endowment* of consumer i. The firm shares of consumer i are represented by a list $\theta^i=(\theta^i_1,\dots,\theta^i_F)$ of non-negative numbers, where θ^i_j is interpreted as the proportion of the total profit of firm j which accrues to consumer i. Hence, the non-labour income of consumer i at the price system p is given by $b^i(p)=pw^i+\Sigma_j\theta^i_j py^j(p)$, and consumer i will choose the consumption plan $x^i(p)=f^i(p,b^i(p))$. By definition of w^i and θ^i, $\Sigma w^i=w$ and $\Sigma_i\theta^i_j=1$ for all $j=1,\dots,F$. (Notice that the model does not explain the ownership relations; these are historically given.)

The model is now fully specified. A *production economy* is given by F firms and H consumers. The firms are described by their technologies $Y_1,\dots,Y_F$; each consumer i is described by his consumption set X_i, his preference relation $\precsim_i$, his initial endowment w^i and his firm shares θ^i, where $i=1,\dots,H$. An *exchange economy* is an economy where no production takes place, i.e, $Y_j=\{0\}$ for all j.

The *total supply* at the price system p is given by $y(p)+w$, where $y(p)=\Sigma y^j(p)$; $x(p)=\Sigma x^i(p)$ is called *total demand*, and $z(p)=x(p)-y(p)-w$ is called *excess demand*. Since all households spend their whole income on consumption, we obtain the so-called *Walras identity* $pz(p)=0$. The price system p^* is said to be an *equilibrium price system* if $z(p^*)=0$. A *market equilibrium* (or *Walrasian equilibrium*) is a collection $\{(x^i(p))_{i=1,\dots,H}, (y^j(p))_{j=1,\dots,F}, p\}$ such that p is an equilibrium price system. (Notice that a proportional change in all prices will not affect the decisions of consumers and producers, i.e., z(p) is homogeneous of degree zero in p.)

Of course, one has to prove that market equilibria exist. One may proceed as follows. Let S be the positive part of the (n-1)-dimensional unit sphere, i.e., $S=\{x\varepsilon R^n: \|x\|=1 \text{ and } x_i>0, i=1,...,n\}$. By Walras' identity and homogeneity of z(p), the excess demand function may be looked at as defining a tangent vector field on S. Consider an exchange economy. If all commodities are desired by the consumers, $\|z(p)\|$ tends to infinity if the price of a commodity approaches zero. Furthermore, each consumer can only supply a certain maximum amount of services to the other members of the society, i.e., the function z(p) is bounded below. These two properties of the excess demand function imply that z(p) points "inward" near the boundary of S (see Figure 1). One can now show that a continuous vector field on S with such a boundary behaviour has at least one point p such that z(p)=0. To prove that an arbitrary continuous function $z: S \longrightarrow R^n$ satisfying

(1) z is bounded below,
(2) $\|z(p)\| \longrightarrow \infty$ if p tends to the boundary of S,
(3) pz(p)=0 for all p in S

must have a zero, one needs Brouwer's fixed point theorem or a similar powerful mathematical argument (see, e.g., Varian, 1984, pp. 195-197; a comprehensive discussion of the existence question is given by Debreu, 1981). The next section will show that it is not possible to establish the existence of a market equilibrium by elementary methods.

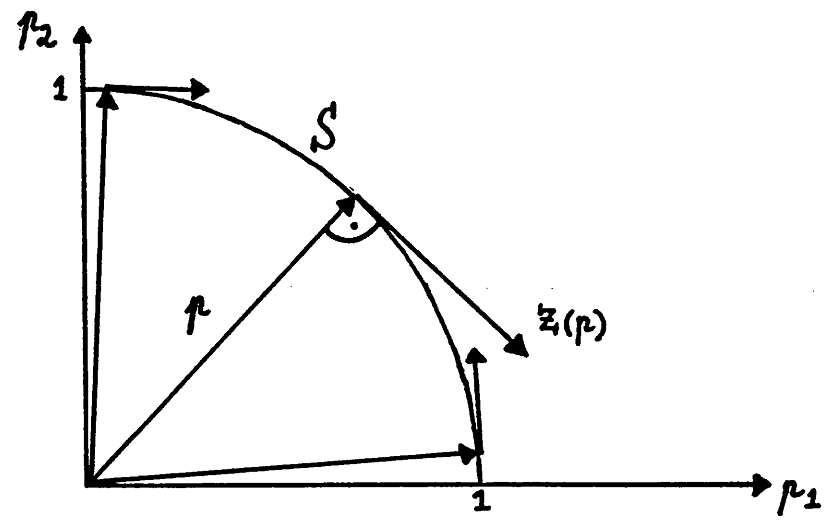


Figure 1

3. Properties of Supply and Demand Functions

The question arises whether microeconomic rationality, i.e., the paradigm of profit and preference maximisation, implies certain structural properties of the excess demand function. We proceed by first proving four simple propositions; we then state two theorems which may be interpreted as the "death sentences" of microeconomic theory. We begin with some general remarks.

3.1. Uniqueness and Stability of Equilibrium

The market equilibrium has been defined as a state of the economy where all consumption and production plans are compatible with each other, i.e., every member of the society can realise his plans. In natural sciences one typically requires much more from an equilibrium. One thinks of the equilibrium of a physical system as an "equilibrium state" and a "motion to the equilibrium state" if the system is not in the equilibrium state already. One may take the view that a complete model of a market economy should have a uniquely determined and (locally) stable equilibrium.

Suppose the economy is in a state of disequilibrium, i.e., the prevailing price system p is such that $z(p)\neq 0$. If p is far away from an equilibrium price system of the economy, it may happen that the economy never reaches an equilibrium state on its own accord since significant disturbances of a system may have unpredictable consequences. However, if p is sufficiently close to an equilibrium price system p^*, and if there are then no forces which guarantee that p tends to p^*, then the Walrasian equilibrium concept is economically not very meaningful. Hence, at the very least, the economy should have a locally stable equilibrium. The problem with multiple equilibria is that it is very difficult to explore how the economy responds to a change in its exogenous parameters (i.e., technologies, preferences, firm shares and resources) if there are a number of distinct equilibria (we will see below that it is not impossible). To put it differently: a model which typically produces more than one equilibrium does not provide a sound theoretical foundation for the great number of comparative-statics studies which one can find in the literature. Let us quote Samuelson (1947, p. 257):

> *Thus, in the simplest case of a partial-equilibrium market for a single commodity, the two independent relations of supply and demand...determine by their intersection the equilibrium quantities of the unknown price and quantity sold. If no more than this could be said, the economist would be truly vulnerable to the gibe that he is only a parrot taught to say "supply and demand." Simply to know that there are efficacious "laws" determining equilibrium tells us nothing of the character of these laws. In order for the analysis to be useful it must provide information concerning the way in which our equilibrium quantities will change as a result of changes in the parameters taken as independent data.*

The problem of stability is more subtle than that of uniqueness. The neoclassical equilibrium model is based on the assumption that all members of the society take prices as given and adjust to them. If at all, however, this behaviour is reasonable only in an equilibrium. The question arises how consumers and producers actually behave in a disequilibrium state. Does this behaviour then imply that a price system p with $z(p) \neq 0$ tends to an equilibrium p^*? In this chapter we will consider a very simple and artificial price adjustment mechanism called the *Walrasian tâtonnement process* (see Arrow and Hahn, 1971, Chapters 11-13, for a detailed discussion of the stability question). The idea is to view the economy as a large "auction". Given a hypothetical price system, consumers and producers inform the auctioneer about their consumption and production plans. The auctioneer, in turn, tries to find prices such that all plans are compatible with each other. This is an artificial but well-defined situation. It is natural to ask whether such an auction will produce an equilibrium. If this is not the case, one may consider more complicated price adjustment processes. However, one should expect that an auction is capable to produce the desired result. Let us be more precise.

Suppose there is a fictitious agent, called the *Walrasian auctioneer*, who announces a price system p. Each producer j informs the auctioneer about his profit maximising production plan $y^j(p)$ at p. The auctioneer informs the consumers about the production plans. Each consumer i uses this information to compute his non-labour income $b^i(p)$ and then informs the auctioneer about his optimal consumption plan $x^i(p)$. The auctioneer computes $z(p)$. If $z(p) \neq 0$, no trade takes place and the auctioneer chooses a new price system such that the price change is proportional to excess demand. The above process begins again and continues until an equilibrium p^* has been found; at p^* all trades will be carried out.

The auction can be described by a system of differential equations $p'_i(t)=k_i z_i(p(t))$ $(k_i>0,\ i=1,\dots,n)$. By choosing the units of measurement appropriately, one may assume without loss of generality that $k_i=1$ for all $i=1,\dots,n$. A very readable book about differential equations is Hirsch and Smale (1974); see especially Ch. 8 and Ch. 9. Let us mention the following:

The excess demand function $z(\cdot)$ is a function from R^n_{++} into R^n. A function $p(\cdot)$: $[0,b] \longrightarrow R^n_{++}$ $(b>0)$ which satisfies $p'(t)=z(p(t))$ is called a *(local) solution* of the differential equation $p'=z(p)$. If $z(\cdot)$ is continuously differentiable, there exists for all p° in R^n_{++} a local solution $p(\cdot)$ of $p'=z(p)$ such that $p(0)=p^\circ$, and $p(\cdot)$ is uniquely determined by the *initial condition* $p(0)=p^\circ$.

Let $p(\cdot)$: $[0,b] \longrightarrow R^n_{++}$ be a solution of $p'=z(p)$. By differentiating $t \longmapsto \|p(t)\|$, one sees that the Walras identity implies $\|p(t)\|$ = constant for all t. Hence, if $p(0) \varepsilon S=\{x \varepsilon R^n: \|x\|=1 \text{ and } x_i>0,\ i=1,\dots,n\}$, the price path $p(\cdot)$ never leaves the unit sphere. By homogeneity of $z(p)$, the excess demand function may be viewed as a function from S into R^n.

We would like to know under what conditions $p(t)$ converges to a zero of $z(\cdot)$ as t tends to infinity and whether the excess demand function satisfies "in general" these conditions. It may, however, happen that there exists no *(global) solution* $p(\cdot)$: $R_+ \longrightarrow S$ with initial value $p(0)=p^\circ$. The reason for this is that there may be a finite $b>0$ such that $p(t)$ tends to the boundary of S as t approaches b. Suppose there is a closed subset $K = K(p^\circ)$ of S, such that every solution $p(\cdot)$: $[0,b] \longrightarrow S$ with $p(0)=p^\circ$ lies entirely in K (i.e., $p(t) \varepsilon K$ for all $t \varepsilon [0,b]$). Then there is a global solution $p(\cdot)$: $R_+ \longrightarrow S$ with $p(0)=p^\circ$ and $p(t) \varepsilon K$, $t>0$. We will assume in the following that this condition is satisfied by the excess demand function (loosely speaking, this means that we exclude the existence of an interval $]0,\bar{p}_i]$ such that $z_i(p)$ is negative for all p_i in $]0,\bar{p}_i]$).

A point x in a subset X of R^n is called *isolated* if x is not a point of accumulation of X, i.e., there exists $\tau>0$ such that $||x-x'||\geq\tau$ for all x' in X with $x'\neq x$. The two crucial definitions are now the following: Let $p(\cdot,p^o)$: $R_+\longrightarrow S$ be the global solution of $p'=z(p)$ with initial condition $p(0)=p^o$. A zero p^* of the excess demand function $z(\cdot)$ is called *locally (Walras) stable* if there exists $\tau>0$ such that $\lim_{t\to\infty}p(t,p^o) = p^*$ for all p^o in S with $||p^*-p^o|| \leq \tau$. The zero p^* of $z(\cdot)$ is called *globally (Walras) stable* if $\lim_{t\to\infty}p(t,p^o) = p^*$ for all p^o in S. Notice that local stability implies that p^* is an isolated zero of $z(\cdot)$ in S while global stability implies that p^* is the only zero of $z(\cdot)$ in S.

Before turning to the properties of demand and supply functions, we want to draw attention to the following point. A commodity is characterised by its *nature* and *quality*, the *place* at which it is available and the *point in time* at which it is available (e.g., oil is a *product* that is available in various qualities; a pianist provides a *service* which is also available in various qualities). An individual may respond to a wage increase by supplying a better quality of labour (see the "efficiency wage" literature, e.g., Akerlof and Yellen, 1986). Likewise, a firm may respond to an increase in the price of one of its products by improving the quality of the product and supplying less units of the improved product.

From the point of view of the model the firm (resp. individual) then no longer supplies the same commodity. By assumption, there is a finite number n of distinguishable commodites in the economy. This means that we have a finite number of exogenously given qualities of products and services, a finite number of exogenously given locations and a finite number of exogenously given periods. If there are K products in the economy, and if the i-th product is available in n_{1i} qualities at n_{2i} places and at n_{3i} points in time, then we have $n = n_{11}n_{21}n_{31} +\ldots+ n_{1K}n_{2K}n_{3K}$ commodities in the economy (see also Malinvaud, 1972, pp. 5-8).

The *price of a commodity* is the amount which has to be paid today for one unit of a product or service of a *given* quality which is available at a *given* location at a *given* point in time. We emphasise that there is no "dishonesty" in the model. If a buyer wants to have x units of a certain commodity i and if the buyer pays today the amount $p_i x$ to the seller, then x units of exactly that commodity will be made available to the buyer. [An excellent discussion of *uncertainty* in the model is given by Debreu (1959, Ch. 7). For a discussion of the problem of *asymmetry of information*, we refer the reader to Varian (1984, Ch. 8); the classic article is that by Akerlof, 1970).

Clearly, in this framework the hypothesis of profit maximisation implies that a firm does not produce less units of a commodity (resp. does not employ more labour) if the price of the commodity (resp. the wage rate) increases while all other prices do not change. More generally, one has the following relation between p and y(p) (note that a function $g:R \longrightarrow R$ is monotone increasing if and only if $(x-x')(g(x)-g(x')) \geq 0$ for all $x,x' \varepsilon R$).

Proposition 1: *Let* y(p) *be the supply function of a profit maximising firm. Then for all price systems* p *and* q

$$\text{(M)} \qquad y(p)=y(q) \text{ or } (p-q)(y(p) - y(q)) > 0.$$

Proof: Suppose $y(p) \neq y(q)$. Let p' be any price system. Since y(p') is the uniquely determined production plan at the price system p', we have $p'y(p') > p'y(p'')$ for all p'' with $y(p') \neq y(p'')$. Hence, $(p-q)(y(p)-y(q)) = [py(p) - py(q)] + [qy(q) - qy(p)] > 0$. Q.E.D.

The property

$$py(p) \geq py(q) \text{ for all } p, q$$

is sometimes called the *weak axiom of profit maximisation*; we will see below that the demand function of a consumer has a similar property. We remark that one may found the theory of the firm on the weak axiom of profit maximisation. Property (M) is additive, i.e., the aggregate supply function $y(p)=\Sigma y^j(p)$ also satisfies (M): Let $y(p) \neq y(q)$. Then there is a firm j such that $y^j(p) \neq y^j(q)$. Hence,

$$(p - q)(y(p) - y(q)) = \Sigma_j (p - q)(y^j(p) - y^j(q)) > 0.$$

Suppose the aggregate demand function $x(p)$ is strictly decreasing on the set S of normalised price systems, i.e., $(p-q)(x(p)-x(q))<0$ for all p, q in S with $p \neq q$. As the next proposition shows, the equilibrium price system is then uniquely determined (up to normalisation) and globally Walras stable.

Proposition 2: *Let* $P \subseteq R^n$ *be open. Let* $z: P \longrightarrow R^n$ *be continuously differentiable. Suppose* (i) $E=\{p \varepsilon P: z(p)=0\}$ *is non-empty and* (ii) *every solution* $p(\cdot)$ *of* $p'=z(p)$ *with* $p(0)=p^o$ $(p^o \varepsilon P)$ *remains in a compact subset of* P. *Then the following holds:*

(a) *if* z *is strictly decreasing on* P, *i.e., for all* p, q *in* P *with* $p \neq q$

$$(p - q)(z(p) - z(q)) < 0,$$

then z *has exactly one zero* p^*, *and* p^* *is globally stable;*

(b) *the zero* p^* *in* E *is isolated and locally stable if there is a* $\tau > 0$ *such that for all* p *in* P *with* $||p-p^*|| \leq \tau$ *and* $p \neq p^*$

$$(p - p^*)z(p) < 0.$$

Proof: Because of (ii) there exists for all p^o in P a uniquely determined function $p(\cdot)$: $R_+ \longrightarrow P$ with $p(0)=p^o$ which satisfies the differential equation $p'=z(p)$ (e.g., Hirsch and Smale, 1974, p. 172).

(a) Obviously, there is at most one p in P such that $z(p)=0$. Let p^* be the unique zero of z. Let $p^o \varepsilon P$, and let $p(\cdot)$ denote the solution of $p'=z(p)$ with initial condition $p(0)=p^o$. If there is a $t^* \geq 0$ such that $p(t^*)=p^*$, then $p(t)=p^*$ for all $t \geq t^*$ since $p(\cdot)$ is uniquely determined by the initial condition $p(0)=p^o$; hence $p(t) \longrightarrow p^*$ as $t \longrightarrow \infty$. Suppose $p(t) \neq p^*$ for all $t \geq 0$. Set $D(t)=||p(t)-p^*||^2$. Then, by strict monotonicity of z,

$$D'(t) = 2(p(t) - p^*)(z(p(t)) - z(p^*)) < 0 \quad (t \geq 0).$$

Hence, $D(t)$ is a strictly decreasing function of t. If $D(t) \longrightarrow 0$ as $t \longrightarrow \infty$, then $p(t) \longrightarrow p^*$. Suppose $D(t) \longrightarrow D$ with $D>0$. We show that this leads to a contradiction: An infinite bounded subset of R^n has at least one point of accumulation. Hence, since $p(t) \varepsilon [D,D(0)]$ $(t \geq 0)$, there exist a sequence (t_n) and a vector p', $p' \neq p^*$, such that $p(t_n) \longrightarrow p'$ as $n \longrightarrow \infty$. Because of (ii), p'

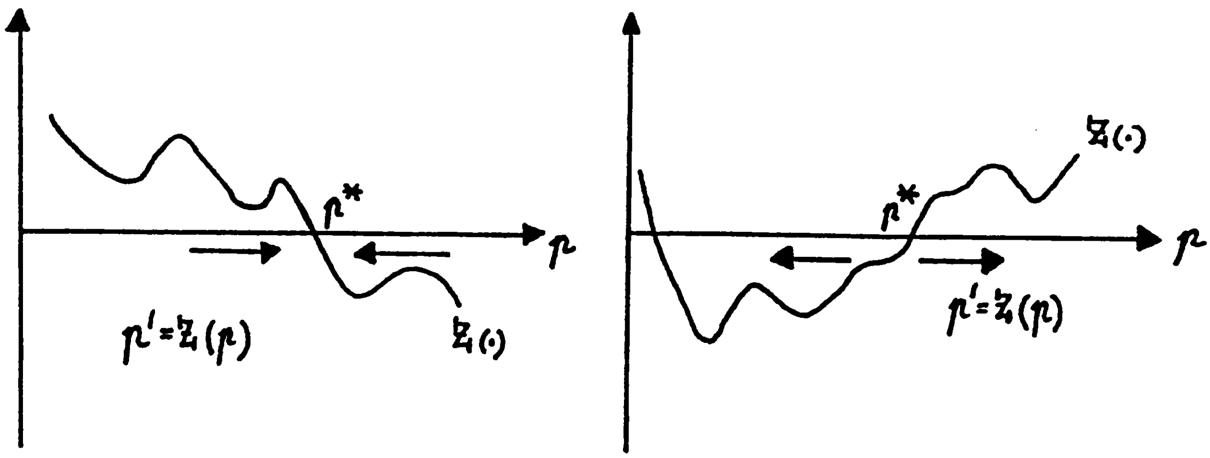
does not lie on the boundary of P, i.e., z(p') is defined. Since $D(t) \longrightarrow D$, we have $D'(t) \longrightarrow 0$ as $t \longrightarrow \infty$. Hence, by continuity of z,

$$0 = \lim_{n \to \infty} D'(t_n) = 2(p' - p^*)(z(p') - z(p^*))$$

However, this is not possible since z is strictly decreasing. Hence, D=0 and therefore $p(t) \longrightarrow p^*$ as $t \longrightarrow \infty$.

(b) Suppose $p(t) \neq p^*$ for all $t \geq 0$. Because of the proof of (a), it suffices to show that $D(0) \leq \tau^2$ implies $D'(t) < 0$ for all $t \geq 0$. But this is obvious: Since $D(0) \leq \tau^2$, we have $D'(0) < 0$, i.e., the distance between p(t) and p* is first decreasing. Suppose there is a $t > 0$ such that $D'(t) \geq 0$. Let t* be the smallest $t > 0$ with $D'(t) \geq 0$. Then $D(t^*) < D(0)$ and $D'(t^*) = 0$. However, $D(t^*) \leq \tau^2$ implies $D'(t^*) = 2(p(t^*) - p^*)z(p(t^*)) < 0$. Hence, $D'(t) < 0$ for all $t \geq 0$. Q.E.D.

Part (b) of Proposition 2 is illustrated in Figure 2a for the case n=1. Notice that (b) does not imply that the partial functions $p_i \longmapsto z_i(p^*_1, \ldots, p_i, \ldots, p^*_n)$ are decreasing on $[p^*_i - \tau, p^*_i + \tau]$. The property of the function z in (b) excludes the instability shown in Figure 2b while (a) implies that the functions $z_i(p)$ $(i=1,\ldots,n)$ are strictly decreasing in their own variable p_i.



(a) (b)

Figure 2

Let us now turn to the implications of the hypothesis of preference maximisation for the market demand function $x(p)$. An individual demand function $f(p,m)$ which is derived from preference maximisation has an important (testable) property:

Proposition 3: *Let* $f(p,m)$ *be the demand function of a preference maximising consumer. Let* p *and* p' *be any two price systems, and let* m *and* m' *be any two non-labour incomes. Then*

$$\text{(W)} \qquad f(p,m) \neq f(p',m') \text{ and } pf(p',m') \leq m \text{ implies } p'f(p,m) > m'.$$

Proof: Suppose $f(p,m) \neq f(p',m')$ and $pf(p',m') \leq m$. Then the consumer could realise the consumption plan $f(p',m')$ at the price system p given non-labour income m. But he chooses $f(p,m) \neq f(p',m')$, and his optimal choice is uniquely determined by p and m. Hence, he prefers $f(p,m)$ to $f(p',m')$. Suppose $p'f(p,m) \leq m'$. Then the same argument produces the contradiction "$f(p',m')$ is preferred to $f(p,m)$". Hence, $p'f(p,m) > m'$. Q.E.D.

Property (W) is called the *weak axiom of revealed preference* (the theory of the consumer may be founded on (W); see Samuelson, 1938). In applications of the equilibrium model one usually assumes that the market demand function $x(p)$ may be looked at as the demand function of a fictitious consumer, called the "representative" consumer, who owns the economy, i.e., the total resources w of the economy are his initial endowment and his firm shares are given by $\theta=(1,\dots,1)$. This assumption does not imply that the partial demand curves $x_i(p)$ $(i=1,\dots,n)$ are decreasing in their own price. Even if the individual demand function $f(p,m)$ is monotone decreasing in p for any given m, the function $p \longmapsto f(p,pw)$, $w \in R^n_+$, may be increasing on a subset of prices (note that an increase in a commodity price p_i implies an increase in income if $w_i > 0$). However, if the market demand function $x(p)$ can be written as $x(p) = f(p,b(p))$, where $b(p) = pw$ + total profit of the production sector at p, then there is "almost always" a uniquely determined and globally stable equilibrium price system.

Proposition 4: *Let* E *be the set of equilibrium price systems of an economy with demand function* x(p)*, supply function* y(p) *and total resources* w*, i.e.,* $E=\{p\varepsilon R^n_{++}: z(p)=0\}$, where $z(p)=x(p)-y(p)-w$. *Let* $\pi(p)=py(p)$*, and let* $N(p)=x(p)-w$. *Suppose the function* x(p) *is the demand function of a fictitious consumer having the initial endowment* w *and profit income* $\pi(p)$. *Then, for any two price systems* p *and* q *in* R^n_{++}

(a) $N(p)\neq N(q)$ *and* $pN(q)\leq\pi(p)$ *implies* $qN(p)>\pi(q)$.

Property (a) *has the following implications*:

(b) $z(p)\neq z(q)$ *and* $pz(q)\leq 0$ *implies* $qz(p)>0$;
(c) $x(p)=x(q)$ *for all* p, q *in* E;
(d) E *is convex, i.e., if* $p,q\varepsilon E$ *then* $\alpha p+(1-\alpha)q\varepsilon E$ $(0\leq\alpha\leq 1)$;
(e) *the equilibrium price system is unique up to a multiplicative constant if* $E\cap S$ *has an isolated point;*
(f) *if* $E\cap S=\{p^*\}$*, then* p^* *is globally Walras stable.*

Proof: By assumption, there exists an individual demand function f(p,m) such that $x(p)=f(p,b(p))$, where $b(p)=pw+\pi(p)$.

(a) "$N(p)\neq Nq)$ and $pN(q)\leq\pi(p)$" is equivalent to "$x(p)\neq x(q)$ and $px(q)\leq b(p)$". Thus, (a) is an immediate consequence of Proposition 3.

(b) Let $z(p)\neq z(q)$ and $pz(q)\leq 0$. Suppose $N(p)=N(q)$. Then $y(p)\neq y(q)$ and therefore $qz(p) = q(N(p)-y(p)) = qN(q)-qy(p) = \pi(q)-qy(p) > 0$. Clearly, $pz(q)\leq 0$ implies $pN(q)\leq\pi(p)$. Hence, if $N(p)\neq N(q)$, then by (a) $qN(p)>\pi(q)$ and therefore $qz(p) = q(N(p)-y(p)) > 0$ since $\pi(q)\geq qy(p)$.

(c) is an immediate consequence of (a) and profit maximisation.

(d) One easily verifies that $\pi(p)$ is convex, i.e., $\pi(\alpha p+(1-\alpha)q) \leq \alpha\pi(p) + (1-\alpha)\pi(q)$ for all p, q $(0\leq\alpha\leq 1)$. Let $p,q\varepsilon E$; put $p_\alpha=\alpha p+(1-\alpha)q$, $0\leq\alpha\leq 1$. Because of (c), $y(p)=y(q)=y$. Since $\pi(p)$ is convex, $\pi(p_\alpha)\leq p_\alpha y$ and therefore $y(p_\alpha)=y$ for all $\alpha\varepsilon[0,1]$. Now $p_\alpha N(p_\alpha) = \pi(p_\alpha) \leq \alpha\pi(p)+(1-\alpha)\pi(q)$. Hence, either $pN(p_\alpha)\leq\pi(p)$ or $qN(p_\alpha)\leq\pi(q)$. Suppose $pN(p_\alpha)\leq\pi(p)$. Since $N(p)=y$ and $\pi(p_\alpha)=p_\alpha y$, we have $p_\alpha N(p)=\pi(p_\alpha)$. Hence, (a) implies $N(p_\alpha)=N(p)$. If $qN(p_\alpha)\leq\pi(q)$, then (a) implies $N(p_\alpha)=N(q)$. Thus, E is convex.

(e) follows immediately from (d).

(f) Let $p(\cdot): R_+ \longrightarrow S$ be a solution of $p'=z(p')$ with $p(t)\neq p^*$ for all $t\geq 0$. Set $D(t)=||p(t)-p^*||^2$. Then $D'(t)=-2p^*z(p(t))<0$ because of Walras' identity and (b). To show $p(t)\longrightarrow p^*$, one now has to repeat the proof of part (a) of Proposition 2. Q.E.D.

The function $N(p)=x(p)-w$ is called *net demand function*. The net demand function $N(p)$ [resp. excess demand function $z(p)$] is said to satisfy the *weak axiom of revealed preference* if (a) [resp. (b)] holds. Notice that (b) does not imply (a); because of the monotonicity of the supply function $y(p)$, the excess demand function $z(p)=N(p)-y(p)$ may satisfy (b) while the net demand function $N(p)$ does not satisfy (a). However, the weak axiom of revealed preference for the market net demand function is the most general condition on the consumption side of the model that by itself (i.e., irrespective of the production sector) guarantees (e) and (f). The following argument is due to H. Scarf (for a detailed discussion see Kehoe, 1985):

Consider an exchange economy, i.e., $pN(p)\equiv 0$. Suppose $N(p)$ does not satisfy (a), i.e., there are price systems p and q with $N(p)\neq N(q)$, $pN(q)\leq 0$ and $qN(p)\leq 0$. Set $Y=\{y\in R^n: py\leq 0 \text{ and } qy\leq 0\}$. Then Y represents a technology with *constant returns to scale*, i.e., $y\in Y$ implies $\alpha y\in Y$ for all $\alpha>0$. Since $N(p)$, $N(q)\in Y$ and $pN(p)=qN(q)=0$, $N(p)$ and $N(q)$ are profit maximising production plans with respect to p and q, respectively. Hence, p and q are equilibrium price systems of the production economy described by $N(\cdot)$ and Y; see Figure 3 on the next page. (Notice that the exchange economy may have a unique equilibrium, i.e., there may be only one p in S such that $N(p)=0$.)

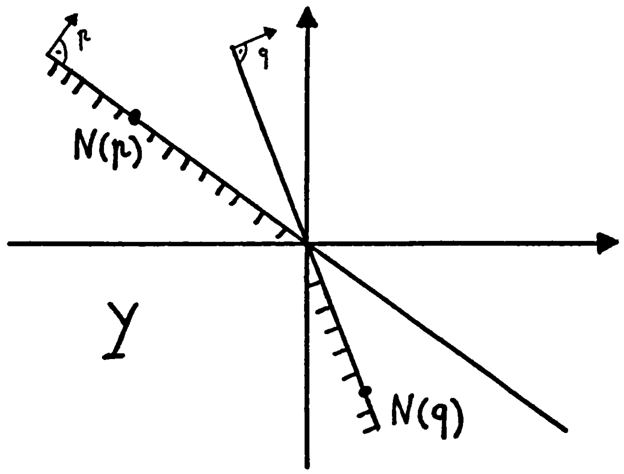


Figure 3

The problem with the weak axiom is that it is not an additive property. One can easily construct examples such that the net demand function N(p) does not satisfy (a). Before turning to the question of what can be said about the class of excess demand functions generated by the equilibrium model, let us explain that it is not restrictive to assume that an equilibrium price system is isolated in S, i.e., "almost all" economies have a finite number of equilibria.

Consider a two-commodity economy with a differentiable excess demand function z(p). By Walras' identity $z_2(p_1,p_2)=-(p_1/p_2)z_1(p_1,p_2)$, and therefore $z(p)=0 \Longleftrightarrow z_1(p)=0$. If one normalises prices such that $p_1^2+p_2^2=1$, then $z_1(p_1,p_2)$ can be represented by a differentiable function $f\colon]0,1[\longrightarrow R$ with $f(x)=z_1(x,\sqrt{1-x^2})$. Let E denote the set of zeros of f. Suppose there are x_1 and x_2 in]0,1[such that $f(x)>0$ on $]0,x_1[$ and $f(x)<0$ on $]x_2,1[$. Then E is closed since f is continuous and the zeros of f stay away from 0 and 1. If $f'(x^*)\neq 0$ at a zero x^*, then one easily verifies that x^* is an isolated zero of f. Hence, if $f'(x)\neq 0$ for all x in E, then E must be finite (see, e.g., Apostol, 1974, Theorem 3.24, p. 54). Furthermore, the boundary behaviour of f implies that f crosses the x-axis at least once and that the number of downward crossings exceeds by exactly one the number of upward crossings, i.e., we obtain an odd number of zeros (see also Figure 4 on page 26).

The above observations can be generalised. An economy is called *regular* if (i) the excess demand function $z\colon R^n_{++} \longrightarrow R^n$ is continuously differentiable and (ii) rank $\delta z(p)=n-1$ at all p with $z(p)=0$ (since $z(\alpha p)=z(p)$, $\alpha>0$, the Jacobian matrix can at most have rank n-1). A regular economy has an odd number of equilibrium price systems in S. Property (ii) is "robust" in the following sense: if at a zero p rank $\delta z(p)<n-1$, one only has slightly to change the excess demand function to obtain rank $\delta z(p)=n-1$; if rank $\delta z(p)=n-1$, then a slight change in z does not lead to rank $\delta z(p)<n-1$ (note: any n linearly dependent vectors in R^n can be made linearly independent by a slight perturbation, but the reverse is not true, i.e., a family of linearly independent vectors cannot be made linearly dependent if only slight variations of the vectors are permitted).

One can show that the excess demand function depends continuously on the parameters of the economy. If a given economy is regular, then it will still be regular after a small perturbation of its parameters has taken place. A slight perturbation of a regular economy does not change the number of equilibria. In particular, one can show that each equilibrium price system moves along a continuously differentiable curve if only local parameter changes are permitted; that is to say, regular economies may be used to study comparative-statics questions. On the other hand, if a given economy is not regular (such an economy is also called "critical"), one only has slightly to change endowments or preferences in order to obtain a regular economy.

The above remarks suggest that regular economies form an open and dense set in the class of all economies generating continuously differentiable excess demand functions. Let us be more specific. Consider an exchange economy where the individual demand functions $f^i(p,m)$ are fixed. Then the parameters of the economy are the initial endowments $w^i \varepsilon R^n$ (i = 1,...,H) of the consumers. One can prove now that those endowment distributions $(w^1,\dots,w^H)$ which give rise to critical economies are contained in a closed and nowhere dense subset of R^{Hn} having Lebesgue measure zero, i.e., they are "negligible" (see Mas-Colell, 1985, for a comprehensive discussion; the classic article is Debreu, 1970). Hence, if one picks randomly an economy then this economy has "almost surely" a finite number of normalised equilibrium price systems. However, unless one makes very strong ad hoc assumptions on agents' characteristics, this is all one can say about the equilibrium states of the model.

3.2. Results on Excess Demand Functions

At the beginnings of the seventies H. Sonnenschein initiated a series of articles on the structure of market demand and excess demand functions (see Shafer and Sonnenschein, 1981, for a survey). It turned out that microeconomic theory does not impose any restrictions on excess demand functions besides continuity, homogeneity, Walras identity and boundary behaviour (i.e., $||z(p)|| \longrightarrow \infty$ as a commodity price tends to zero).

In the following we consider exchange economies such that all consumers have continuous, strictly convex and monotone preference relations (see Debreu, 1959, Chapter 4, for definitions), i.e., the underlying exchange economy does not have "pathological" features. Let $z^i(p)$ be the excess demand (i.e., demand minus endowment) of consumer i at the price system p.

Given any compact set K of price systems, the question arises what can be said about the class of excess demand functions generated by exchange economies on K. By homogeneity of the excess demand function we may assume that K is a subset of S. Let $S_\varepsilon = \{p \varepsilon S: p_i \geq \varepsilon, i=1,\dots,n\}$ $(\varepsilon > 0)$.

In an impressive paper Sonnenschein (1973) showed that the class of excess demand functions, as functions from S_ε into R^n, lies dense in the set of all continuous functions $z: S_\varepsilon \longrightarrow R^n$ with $pz(p)=0$ for all p in S_ε. Sonnenschein proceeded as follows: He fixed the price of the n-th commodity at unity and considered an arbitrary compact set K of strictly positive price systems $p = (p_1,\dots,p_{n-1})$ [this specific price normalisation is inessential for the result]. He then showed that, given any polynomials $z_i: R^{n-1} \longrightarrow R$ $(i=1,\dots,n-1)$, there is an exchange economy whose excess demand on K is equal to $(z_1(p),\dots,z_{n-1}(p),z_n(p))$, where (by Walras' identity) $z_n(p) = -p_1 z_1(p)-\dots-p_{n-1} z_{n-1}(p)$. Some months later the "state of the art" result was published.

Theorem (Debreu, 1974): *Let* $z: S \longrightarrow R^n$ *be a continuous function with* $pz(p)=0$ *for all* p *in* S. *Let* $\varepsilon > 0$. *Then there exists an exchange economy with* n *consumers such that* $\Sigma z^i(p) = z(p)$ *for all* p *in* S_ε.

The theorem tells us that, apart from boundary behaviour, any tangent vector field defined on the positive portion of the (n-1)-dimensional unit sphere $\{x \in R^n: ||x|| = 1\}$ may be looked at as the excess demand function of an exchange economy. One therefore should not expect that a given Walrasian equilibrium is stable with respect to the adjustment process $p' = z(p)$. In fact, if $n \geq 3$, the differential equation $p' = z(p)$ may have very complex solutions.

It is worth mentioning that one obtains a misleading picture of the situation if one studies the question of stability using a two-commodity economy. In this case an equilibrium price system may be *locally* unstable. However, the adjustment mechanism $p' = z(p)$ takes the economy always to *some* equilibrium price system. Figure 4 below illustrates this. As we see, the equilibria A, C and E are locally stable; the equilibria B and D are unstable.

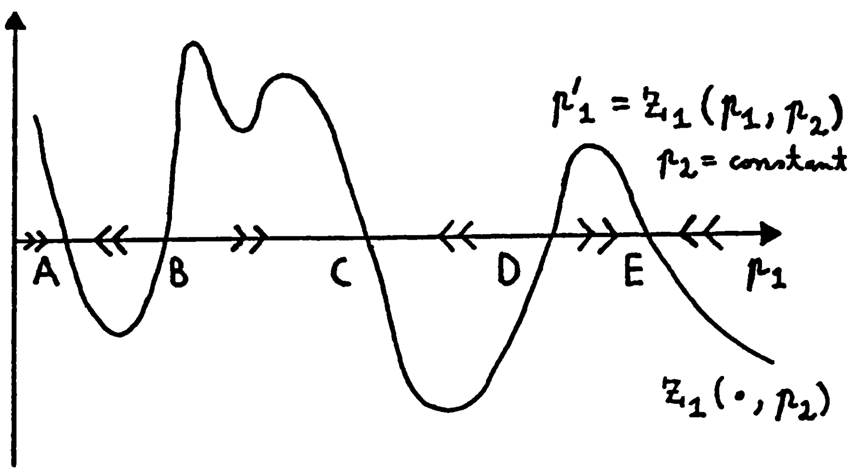


Figure 4

We remark that already Scarf (1960) showed that an exchange economy may have an unstable equilibrium. More precisely, Scarf gave an example of an economy with three commodities and three consumers which has a uniquely determined equilibrium, but no solution of $p' = z(p)$ converges to the equilibrium price system. Later Gale (1963) gave an example of an unstable three-commodity economy with only two agents.

The reader will see in the next section how Debreu proved his theorem; a closer look at his proof reveals that even fixing the distribution of initial endowments does not give more structure. To see why Debreu needs at least as many consumers as commodities, consider the case n=2 and a one-consumer economy. Then the excess demand function satisfies the weak axiom of revealed preference and hence the theorem fails to hold. One may, however, assume that there are more consumers than commodities (Section 4).

In fact, it follows from Debreu's proof that the given excess demand function may be produced by a group of H consumers, where H is any integer which is greater than or equal to n, having identical preferences, i.e., $\precsim_1 = \ldots = \precsim_H$, and collinear endowments, i.e., the endowment bundles of the H consumers lie on a straight line (Kirman and Koch, 1986).

This lack of structure is perhaps less amazing than it appears at first glance. Firstly, the weak axiom of revealed preference, satisfied by an individual demand function, is not an additive property. Secondly, to construct an economy with the desired properties, one may arbitrarily pick some points in the extremely large set of agents' characteristics (i.e., preferences and endowments). Certainly, the consumption sector of a real economy exhibits considerably more structure than the consumption sector of the model. Thirdly, the price dependence of non-labour income may "wipe out" properties of an individual demand functions f(p,m).

The last observation was already utilised by Scarf (1960). Sixteen years later R. Mantel published a striking theorem which we will state below and which is based on exactly this observation. Suppose the preference relation of a consumer has the following property: if $x \precsim x'$ then $\alpha x \precsim \alpha x'$ for all $\alpha > 0$. Such preferences are called *homothetic* and imply that the demand function f(p,m) can be written as f(p,m)=g(p)m (we remark that the three consumers in Scarf's example have homothetic preferences). Before stating Mantel's result, we show that the weak axiom implies that the function g(p) is monotone decreasing.

Proposition 5: *Let* f(p,m)=g(p)m. *Suppose* f(p,m) *satisfies* (W) *and the Walras identity, i.e.,* pf(p,m)=m. *Then for all price systems* p *and* q

$$(p - q)(g(p) - g(q)) \leq 0.$$

Proof: Let $f(p,m)$ be an arbitrary individual demand function. Let $m' = qf(p,m)$. Then the weak axiom implies $(p-q)(f(p,m)-f(q,m')) \leq 0$ (the inequality represents the substitution effect of a price change from p to q; the non-labour income m of the consumer has been compensated in such a way that he can realise $f(p,m)$ at the price system q). Let $m=1$ and $f(p,m)=g(p)m$. Then

$$
\begin{aligned}
(p - q)(g(p) - g(q)) &= (p - q)(g(p) - g(q)m') + (p - q)(g(q)m' - g(q)) \\
&\leq (p - q)g(q)(m' - 1) \\
&= (pg(q) - 1)(qg(p) - 1) \quad \text{(since } pf(p,m)=m\text{)}.
\end{aligned}
$$

Suppose $g(p) \neq g(q)$ and $pg(q) \leq 1$ [resp. $qg(p) \leq 1$]. Then the weak axiom implies $qg(p) > 1$ [resp. $pg(q) > 1$]. Suppose $pg(q) > 1$ and $qg(p) > 1$. Then $(p-q)(g(p)-g(q)) = [1-pg(q)] + [1-qg(p)] < 0$. This completes the proof.

Hence, if the individual demand functions $f^i(p,m)$ $(i=1,\ldots,H)$ are of the form $f^i(p,m)=g^i(p)m$ and if the non-labour income of the consumers does not depend on p, then the market demand function $x(p) = g^1(p)b^1+\ldots+g^H(p)b^H$ (b^i denoting the non-labour income of consumer i) is monotone decreasing in p. However, in an exchange economy the non-labour income of the consumers is given by $b^i=pw^i$, $w^i \in R^n_+$ $(i=1,\ldots,H)$. One immediately verifies that the scalar product $(p - q)\cdot(g^i(p)\cdot pw^i - g^i(q)\cdot qw^i)$ can be written as

$$pw^i\cdot(p - q)\cdot(g^i(p) - g^i(q)) + \{(p - q)\cdot g^i(q)\}\cdot\{w^i\cdot(p - q)\}.$$

Notice that the second term represents the income effect of a price change from p to q. It turns out that the income functions $b^i(p)=pw^i$ $(i=1,\ldots,H)$ may completely "wipe out" the monotonicity of the function $q \longmapsto \Sigma\alpha_i g^i(q)$, where $\alpha_i=pw^i$ $(i=1,\ldots,H)$.

Theorem (Mantel, 1976): *Let* $z\colon S \longrightarrow R^n$ *be a function with continuous second-order partial derivatives and* $pz(p)=0$ *on* S. *Let* $w^1,\ldots,w^n$ *be any linearly independent vectors in* R^n_+. *Let* $\varepsilon>0$. *Then there exist a constant* $k>0$ *and* n *homothetic preference relations* $\precsim_i$ $(i=1,\ldots,n)$ *so that the exchange economy* $(\precsim_i, kw^i)_{i=1,\ldots,n}$ *generates* z *on* S_ε, *i.e.,*

$$\Sigma z^i(p) = z(p) \text{ for all p in } S_\varepsilon.$$

We remark that there are two special cases where homothetic preferences imply that the exchange economy behaves as if a single consumer were maximising homothetic preferences: (i) If the consumers have identical homothetic preferences, i.e., $f^i(p,m)=g(p)m$ for all $i=1,\dots,H$, and arbitrary initial endowments w^i, then the market demand function $x(p)$ is given by $x(p)=g(p)\cdot pw$, where $w=\Sigma w^i$. (ii) Suppose the functions $g^i(p)m$ are arbitrary and the initial endowments w^i are collinear, i.e., there are a vector $w\varepsilon R^n_+$ and numbers $\alpha_i>0$ such that $w^i=\alpha_i w$, $i=1,\dots,H$. Then market demand becomes $x(p)=g(p)\cdot pw$, where $g(p)=\Sigma\alpha_i g^i(p)$. One can prove that the function $f(p,m)=g(p)m$ is the demand function of a consumer having homothetic preferences (see, e.g., Shafer and Sonnenschein, 1981, Theorem 3, p. 676).

However, Mantel's theorem tells us that a slight deviation from proportional initial endowments is sufficient to obtain market excess demand functions which may have any structure [note that if all preference relations are homothetic and the exchange economy $(\preccurlyeq_i, kw^i)_{i=1,\dots,H}$ $(k>0)$ generates the excess demand function $z(p)$, then the exchange economy $(\preccurlyeq_i, w^i)_{i=1,\dots,H}$ generates the excess demand function $(1/k)z(p)$].

The results of Sonnenschein, Debreu and Mantel show that strong assumptions are required if one wants to use the equilibrium model in order to study problems which go beyond the "three famous questions" of economics, i.e., existence of market equilibria, Pareto optimality of market equilibria and decentralisation of Pareto optimal states. If one picks randomly an (regular) exchange economy then the excess demand function of this economy may have very "peculiar" features. To put it differently, those exchange economies which generate macroeconomic regularities form a small set in the class of all exchange economies. We remark that Hildenbrand (1989b) has recently shown that the market net demand function "always never" satisfies the weak axiom of revealed preference.

Recall that the aggregate supply function is monotone decreasing in the commodity prices (Proposition 1). Thus, if there are multiple and unstable equilibria in a production economy, then the source of instability and multiplicity lies entirely in the consumption sector of the model.

The question arises whether there are "reasonable" assumptions on the distribution of consumers' characteristics which imply that the market demand function has certain structural properties. To get a feel for this question, let us return to the equilibrium concept of the model.

The model is composed of a finite number of consumers and a finite number of producers. For the *proof* that an equilibrium exists, the number of consumers (resp. producers) is inessential. However, the *notion* of a competitive equilibrium makes economic sense only if there is a large number of traders such that each individual trader has no influence on the state of the economy. Certainly, the members of a small community will not decentralise their decisions via a price system. A trader takes prices as given and adjusts to them only if he is of the opinion that he cannot change the prevailing prices.

It seems natural to model the consumption side of the economy as a very large population of individually small and "price-taking" persons. Aumann (1964) suggested that the appropriate mathematical model for such a population "is one in which there is a continuum of traders (like the continuum of points on a line)" (p. 39). It should be intuitively clear that a given economy with a continuum of individuals may be viewed as the "limit" of a sequence of finite economies having the property that the number of individuals tends to infinity (details can be found in Hildenbrand, 1974).

Of course, the production sector of a modern economy is not just a collection of technologies. In the model the only difference between production and consumption is that the assumption of profit maximisation immediately implies that the supply function is a decreasing function of the commodity prices. In a real economy, however, a very large number of individually small consumers faces firms which vary in their size. At least the big firms do not buy and sell commodities at an exogenously given price structure.

Note that one has to define a new equilibrium concept if there are "price-making" firms in the economy. One then has to study the uniqueness and "stability" of this new equilibrium concept. We remark that the *theory of the firm* (or, more generally, the *theory of industrial organisation*) has become a major research area. A useful textbook is Tirole (1988); extensive surveys of various aspects of business behaviour can be found in the Handbook of Industrial Organization (Schmalensee and Willig, 1989).

The following three chapters will be concerned with the consumption side of the economy. In the next chapter we will leave the framework of the neoclassical equilibrium model. We will consider a very simple consumption sector. All individuals have identical preferences but they differ in their income. The two essential features of the model are the following. Firstly, identical individuals earn different wage rates. Secondly, wage rates are continuously distributed in a finite interval [a,b], i.e., the population consists of a continuum of individuals.

Our focus of attention will be the dependence of the market demand function upon the distribution of wages in the population. Since we face a continuum of individuals, summation will be replaced by integration.

As in this chapter, we will consider in the next chapter "rational" consumers. More precisely, the individual consumption behaviour will be represented by a demand function which satisfies a weak version of the weak axiom of revealed preference. One should view this assumption simply as a "working hypothesis". On the micro-level there may be very complex socio-economic interactive processes. However, the market demand function may satisfy the weak axiom of revealed preference even if no individual demand function satisfies the axiom.

It should be mentioned that psychologists have carried out a great number of experiments. The results of these experiments suggest that people do not behave "rationally", i.e., the departures of the *actual* behaviour of people from *rational economic* behaviour appear to be of a systematic nature. For a discussion of this experimental evidence and its relevance for economics we refer to the articles collected in Hogarth and Reder (1987). In particular, we refer the reader to Herbert Simon's fundamental contributions on the problem of rationality; see Simon (1955, 1956, 1972, 1976). In fact, Simon argues that people do not *maximise* but that they merely *"satisfice"*, that is to say, an individual tries to realise a given "aspiration level", but the *actual* decision of a person is typically not the *best* decision in the huge set of all *conceivable* decisions.

Ideally, one should therefore try to derive properties of the market demand function from assumptions on the distribution of consumers' characteristics without postulating economic rationality on the micro-level. We remark that such results were recently obtained by Grandmont (1992).

At the end of this chapter we want to show what happens if one fixes in Debreu's (1974) theorem not only an excess demand function $z: S \longrightarrow R^n$ but also the collection $(w^1,\dots,w^H)$ of individual endowment bundles. Recall that we say that an exchange economy $(\precsim_i, w^i)_{i=1,\dots,H}$ *generates* $z(\cdot)$ on $S_\varepsilon=\{x \varepsilon R^n: ||x||=1 \text{ and } x_i \geq \varepsilon, i=1,\dots,n\}$ $(\varepsilon>0)$ if the individual excess demand functions $z^i(\cdot)$ $(i=1,\dots,H)$ sum up to $z(\cdot)$ on S_ε, i.e., $\Sigma z^i(p) = z(p)$ for all p in S_ε.

4. Debreu's Theorem

Fixing the initial endowment bundles w^i restricts the class of excess demand functions z(p) since market demand x(p) is non-negative if consumers do not supply labour and $x(p)=z(p)+\Sigma w^i$. However, one does not obtain more "structure" if one fixes $(w^1,\dots,w^H)$. In view of Mantel's result the following observation is not surprising (see also Kirman and Koch, 1986). The reader will now see how Debreu proceeded to prove his theorem.

Claim: *Let* $z: S \longrightarrow R^n$ *be continuous with* $pz(p)=0$ *for all* p *in* S. *Let* $w^1,\dots,w^H$ *be any vectors in* R^n_{++}. *Let* $H \geq n$ *and* $\varepsilon>0$. *Then the following holds:*

(a) *there exist a constant* $k>0$ *and preference relations* $\precsim_i$ $(i=1,\dots,H)$ *such that the exchange economy* $(\precsim_i, kw^i)_{i=1,\dots,H}$ *generates* z *on* S_ε;
(b) *there exist a constant* $k>0$ *and preference relations* $\precsim_i$ $(i=1,\dots,H)$ *such that the exchange economy* $(\precsim_i, w^i)_{i=1,\dots,H}$ *generates* $k\cdot z$ *on* S_ε.

The $\precsim_i$ $(i=1,\dots,H)$ *may be chosen so that each* $\precsim_i$ *is a continuous, strictly convex and monotone preference relation on* R^n_+.

Proof: Debreu constructs the individual excess demand functions $z^i: S \longrightarrow R^n$ which sum up to z on S_ε as follows: Let p be in S_ε. Set $a(p)=z(p)+\Theta(p)\cdot p$, $\Theta(p)\varepsilon R$, and choose $\Theta(p)$ large enough so that $a(p)>>0$ (see Figure 5). Since z is continuous, we can choose $\theta(p)$ so that the function $\theta\colon S_\varepsilon \longrightarrow R$ is continuous.

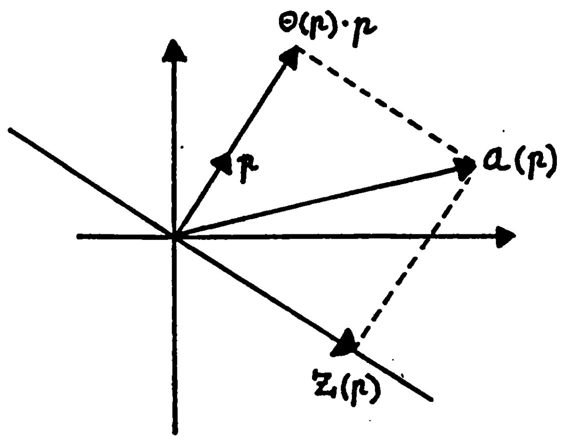


Figure 5

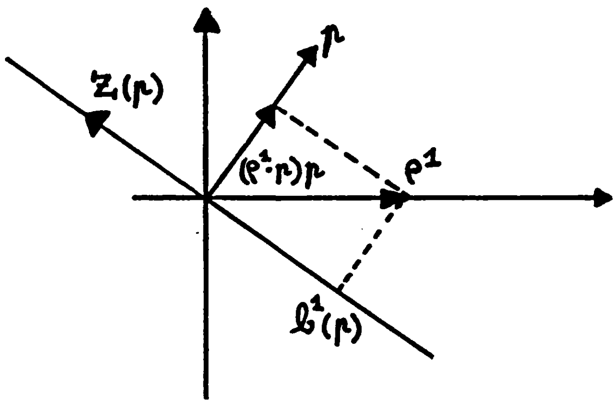


Figure 6

Set $b^i(p)=e^i-p_i\cdot p$ $(i=1,\ldots,n)$, where e^i denotes the i-th unit vector in R^n; i.e., $b^i(p)$ is the orthogonal projection of e^i onto the one-dimensional subspace generated by $z(p)$ (see Figure 6 above). Finally, set $z^i(p)= a_i(p)b^i(p)$ $(i=1,\ldots,n)$, where $a_i(p)$ denotes the i-th component of $a(p)$. One immediately verifies that the $z^i(p)$ satisfy Walras' identity and sum up to $z(p)$. It is not difficult to show that the $z^i(p)$ satisfy the *strong axiom of revealed preference* which implies that each $z^i(p)$ is the excess demand function of a "fully rational" consumer (here we need that $a_i(p)$ is greater than zero; for the definition of the strong axiom of revealed preference see, e.g., Varian, 1984, p. 143; see also Shafer and Sonnenschein, 1981, p. 680). The difficult part of Debreu's proof is to show that the $z^i(p)$ can be generated by preference relations which have no "pathological" features.

Since $z^i(p)$ is continuous and S_ε is compact, there exists $w^i \varepsilon R^n_+$ such that $z^i(p)+w^i \gg 0$ on S_ε. If we were free to choose the individual endowment bundles, we could pick now any vectors $w^1,\ldots,w^n$ such that $z^i(p)+w^i \gg 0$ on S_ε $(i=1,\ldots,n)$; see Figure 7. For prices p in S_ε the demand $x^i(p)$ of consumer i is then given by $x^i(p) = z^i(p) + w^i$.

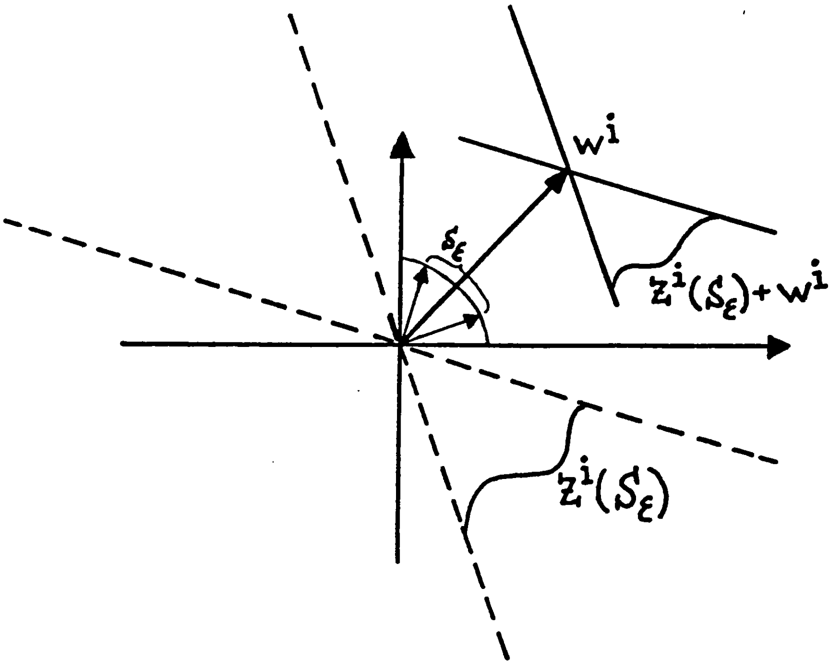


Figure 7

We now turn to (a) and (b):

(a) Since $w^i \gg 0$, there exists $k_i > 0$ such that $z^i(p)+k_i w^i \gg 0$ for all p in S_ε $(i=1,\dots,n)$. Set $k = \max\{k_1,\dots,k_n\}$. Then $z^i(p)+kw^i \gg 0$ on S_ε $(i=1,\dots,n)$. Hence, there are preference relations $\precsim_i$ $(i=1,\dots,n)$ such that the exchange economy $(\precsim_i, kw^i)_{i=1,\dots,n}$ generates z on S_ε.

(b) Let k be as in (a). Set $k'=1/k$. Then $k'z^i(p)+w^i \gg 0$ for all p in S_ε and all $i=1,\dots,n$. By definition of $z^i(p)$, $k'z^i(p) = (k'z_i(p)+k'\theta(p)\cdot p_i)\cdot b^i(p)$. Hence, there are preference relations $\precsim_i$ $(i=1,\dots,n)$ such that the exchange economy $(\precsim_i, w^i)_{i=1,\dots,n}$ generates $k'z$ on S_ε.

If one wants to have more consumers than commodities, i.e., $H>n$, one may proceed as follows. Let $z^i(p)$ $(i=1,\dots,n)$ be as defined above. Set $K=H+1-n$. Pick consumer n and set $\tilde{z}^j(p)=(1/K)z^n(p)$, $j=n,\dots,H$; i.e., the excess demand of consumer n is viewed as the sum of the excess demands of K consumers all of them having 1/K of the excess demand of consumer n. Now apply the above arguments to the excess demand functions $z^1(p),\dots,z^{n-1}(p)$, $\tilde{z}^n(p),\dots,\tilde{z}^H(p)$. This completes the proof.

Chapter 1

Labour Supply Functions, Wage Rate Distributions and the "Law of Demand"

1. Introduction

Let us consider a population of households who have identical preferences but differ in their income. Such a consumption sector can be described by a function f from $R^n_{++} \times R_+$ into R^n_+ and a probability measure μ on R_+. Here f(p,m) is interpreted as the commodity bundle demanded at the price system p by a household with income m, and μ represents the income distribution in the population. The market demand function, F, is then obtained by integrating $f(p,\cdot)$ with respect to μ. It will be assumed that each household spends all its income on consumption, i.e., households can not keep money back.

The market for commodity i is said to fulfill the "law of demand" if the partial demand function $p_i \mapsto F_i(\bar{p}_1,\ldots,p_i,\ldots,\bar{p}_n)$ is monotone decreasing for any given prices $\bar{p}_j$, $j \neq i$. The following property is a generalisation of the "law of demand": the function F is called *monotone (decreasing)* if for any two different price vectors p and q the vectors p-q and F(p)-F(q) point in "different" directions, i.e., the angle between them is not smaller than 90°; if the angle is always greater than 90°, then the function F is called *strictly monotone*.

Notice that the price independence of the income distribution implies that F is not homogeneous of degree zero in p. In Chapter 0 we have seen that a strictly monotone demand function implies the uniqueness and stability of the market equilibrium. More precisely, if $y \in R^n_+$ is an exogenously given supply vector and p^* is such that $F(p^*)=y$, then p^* is unique and every solution of $p'=F(p)-y$ converges to p^* (Proposition 2, page 18).

It is usually assumed that the individual demand function f is derived from utility maximisation. However, the utility hypothesis alone does not give structure in the aggregate: the excess demand function in an exchange economy with identical individuals may have any structure (see Kirman and Koch, 1986). A weaker requirement on individual rationality than the utility hypothesis is the so-called ***weak axiom of revealed preference***. The axiom implies that the market demand function F is monotone if the matrix of *mean income effects*, which one obtains by integrating $f_i(p,\cdot)\delta_m f_j(p,\cdot)$ with respect to μ for all i,j, is positive semi-definite at each point p.

The importance of the income distribution for the structural properties of the function F was emphasised by Hildenbrand (1983). He shows that for every continuous individual demand function f which satisfies a weak version of the weak axiom of revealed preference the market demand function is monotone if the distribution μ can be represented by a decreasing density function on R_+; if, in addition, the density function is concentrated on a finite interval [0,b] and f is a continuously differentiable function which satisfies a weak regularity assumption (i.e., the rank of the substitution matrix of f has to be equal to n-1 for all prices p and incomes m), then the market demand function is even strictly monotone.

In this chapter we incorporate into the above model of a homogeneous household sector an individual labour supply function. Hence, the income distribution is now no longer exogenous. Its shape will depend on the individual labour supply behaviour, and the distribution will also depend on the consumer goods prices.

The chapter is organised as follows. The next section introduces some definitions and sketches the proof of Hildenbrand's theorem. Section 3 extends the model. It is straightforward to see that in the extended model integrating over non-labour income yields aggregate demand functions which are in general not monotone irrespective of the distribution of non-labour income. In the remaining part of the section the assumption that a given commodity has exactly one price will be dropped for the labour market, i.e., we will integrate the individual demand function with respect to a distribution of wage rates. Section 4 contains some final remarks.

It should be emphasised that all individuals of the given group supply the same type of labour. Postulating that their wage rates differ means therefore that we implicitly assume that the labour market does not function in the same manner as the consumer goods markets.[1]

We will not be able to say very much about the aggregate labour supply function. However, the individual labour supply function allows a closer look at the income distribution. In Hildenbrand's analysis the price independence of the income distribution plays an important role. If the income distribution depends on prices then one can construct examples such that the market demand function is not monotone, even if for each price vector p the corresponding income density is decreasing on R_+. On the other hand, individual labour supply is a function of the wage rate *and* the consumer goods prices. In our approach, however, it is not the price dependence of labour supply but only the price dependence of the wage rate distribution that may cause problems. This suggests that future research should explore the dependence of the wage rate distribution on commodity prices.

2. Hildenbrand's Approach

In the following P denotes the set of all strictly positive vectors of R^n, where n represents the number of commodities; a generic element of P is denoted by p and is interpreted as a price system. Individual income is represented by $m \in R_+$.

The function $f: P \times R_+ \longrightarrow R^n$ is said to fulfill the **(weak version of the) weak axiom of revealed preference** if $p'f(p,m) \leq m'$ implies $pf(p',m') \geq m$ for all (p,m) and (p',m') in $P \times R_+$.[2] We say that f satisfies the **Walras (or budget) identity** if $pf(p,m)=m$ for all (p,m). A continuous function f from $P \times R_+$ into R^n_+ which satisfies the weak axiom of revealed preference and the budget identity is called an **individual demand function.**

Notice that a function which fulfills the weak axiom of revealed preference and the Walras identity may have component functions which take on negative values. However, in this section an individual demand function is assumed to have non-negative component functions, i.e., the behaviour described by f does not include labour supply.

Let μ be a probability measure on R_+. Then the non-negativity of an individual demand function together with Walras' identity implies that $f(p,\cdot)$ is μ-integrable, i.e., $\int f_h(p,\cdot)d\mu<\infty$ for all component functions f_h, if and only if μ has a finite mean. A (price independent) probability measure μ on R_+ with finite mean is called an **income distribution.**

Let μ be an income distribution and f be an individual demand function. Then the function $F: P \longrightarrow R^n_+$ defined by $F(p) = \int f(p,\cdot)d\mu$ is called **per capita (mean, aggregate) demand function.** If μ has a density ρ, then $F(p) = \int f(p,m)\rho(m)dm$. The density ρ is said to be **decreasing** on R_+ if $\rho(m_1) \leq \rho(m_2)$ for all $m_1 \geq m_2 \geq 0$.

We remark that f(p,m) may be interpreted as the mean demand at the price system p of all households with income m in a heterogeneous population (i.e., people differ also with respect to their consumption behaviour). In this case, however, the function f does not necessarily satisfy the weak axiom of revealed preference. Whether or not the axiom is fulfilled on the aggregate level will depend on the "form" of the joint distribution of income and tastes; see Hildenbrand (1985b, p. 45) and Hildenbrand (1989a, Proposition 2, p. 271).

Theorem (Hildenbrand, 1983, p. 1003): *Let* f *be an individual demand function and* μ *be an income distribution which can be represented by a decreasing density function on* R_+. *Then the per capita demand function,* F, *is monotone, i.e., for all* p, q *in* P *we have*

$$(p - q)(F(p) - F(q)) \leq 0.$$

It is crucial for the subsequent section to understand why the theorem is true. We will therefore sketch the proof:

Since μ has a decreasing density, the so-called *second mean value theorem for integrals* [3] implies that F is monotone if and only if for all b>0 the function

$$p \longmapsto \int_0^b f(p,m)\,dm$$

is monotone. This is the easy part of the proof and allows Hildenbrand to restrict attention to uniform income distributions. Suppose the given individual demand function f is continuously differentiable and homogeneous of degree zero, i.e., $f(\tau p,\tau m)=f(p,m)$ for all $\tau>0$. Then the remaining part of the proof rests on the *Slutzky equation*, which says that for all i and j

$$\delta_{p_j} f_i(p,m) = \delta_{q_j} s_i(q,m)|_{q=p} - f_j(p,m)\cdot\delta_m f_i(p,m),$$

where, for fixed p and m, the map $q \longmapsto s(q,m)$ is the *Slutzky compensated demand function* corresponding to f(p,m), i.e., $s(q,m)=f(q,qf(p,m))$ for all q in P. The first term on the right-hand side represents the *substitution effect* resulting from a price change and the second term represents the *income effect*. The proof proceeds now in three steps:

(1) A continuously differentiable homogeneous function f(p,m) with $pf(p,m)= m$ satisfies the weak axiom if and only if the matrix of substitution effects of f(p,m) is negative semi-definite for all (p,m); see Kihlstrom et al., 1976, Theorem 1 and 3, pp. 974-975 (we remark that their Condition 4 is equivalent to the definition of the weak axiom given above).

(2) Integrating the Slutzky equation yields

$$\delta_{p_j} F_i(p) = \frac{1}{b}\cdot\int_0^b \delta_{q_j} s_i(q,m)|_{q=p} dm - \frac{1}{b}\cdot\int_0^b f_j(p,m)\cdot\delta_m f_i(p,m)\,dm,$$

where the first term on the right-hand side is the *per capita substitution effect* of a price change and the second term is the *per capita income effect* (with respect to a uniform income distribution).

(3) F is monotone if and only if the Jacobian matrix of F is negative semi-definite for all prices p (see, e.g., Ortega and Rheinbold, 1970, pp. 141-142). Because of (1), the matrix of per capita substitution effects is negative semi-definite for any given income distribution. The matrix of per capita income effects is in general not positive semi-definite [which would imply negative semi-definiteness of the Jacobian matrix $\delta F(p)$]. However, if the distribution of income is uniform on [0,b], then the income effect matrix is positive semi-definite (we will see in the next section that here the non-negativity of f is crucial).

Thus, given a uniform income distribution, the Jacobian matrix $\delta F(p)$ is negative semi-definite for all p in P. Since the theorem holds for a smooth individual demand function, we will expect it to be valid for any continuous demand function. The proof of the general case, however, requires hard work. Hildenbrand does not assume that the individual demand function f is homogeneous of degree zero. Thus, in general, there exists no sequence (f_n) of homogeneous smooth demand functions such that, for given p, $f_n(p,\cdot)$ converges uniformly to $f(p,\cdot)$ on [0,b] (which would imply that the function $F(p) = \frac{1}{b}\cdot\int_0^b f(p,m)dm$ is monotone).

It is natural to ask whether the theorem would also be valid with another type of income density. A closer look at the proof, however, reveals that this is not case.[4] We remark that the graph of the function $f_i(p,\cdot)$ is called the *Engel curve* for commodity i (at the price system p).

Proposition 1: *Let* ρ *be a continuously differentiable density function. Suppose* ρ *is concentrated on a finite interval* [0,b] *and strictly increasing somewhere. Then there exists a continuously differentiable individual demand function* f *such that the aggregate demand function* F *is not monotone. The function* f *may be chosen so that it is derived from utility maximisation and has increasing Engel curves on* R_+ *(i.e.,* $\delta_m f_i(p,m)\geq 0$ *for all* $m\geq 0$ *and all* $i=1,\dots,n$).

Notice that by the Slutzky decomposition monotone increasing Engel curves imply that the partial demand functions F_i are decreasing in their own price irrespective of the income distribution. We will prove Proposition 1 at the end of the next section. We remark that one can use the same method of proof in order to show that if the income density is not decreasing on R_+, then there exists an individual demand function f such that the market demand function F does not satisfy the weak axiom of revealed preference, i.e., there are price vectors p and q with $F(p)\neq F(q)$ so that $pF(q)$ and $qF(p)$ are not greater than the mean of the income distribution; see Freixas and Mas-Colell (1987, Proposition 1, p. 520).

3. Labour Supply and Commodity Demand Functions

In order to incorporate a labour supply function into the model, we have to slightly change the definition of an individual demand function. We will consider continuously differentiable demand functions which are homogeneous of degree zero.

Definition: A continuously differentiable function $f: P \times R_+ \longrightarrow R^{n+1}$, where P denotes the set of all positive vectors of R^{n+1}, is called *individual demand function* if

(a) f satisfies the weak axiom of revealed preference and the Walras identity (see Section 2), and f is homogeneous of degree zero;

(b) the component functions $f_1,\dots,f_n$ are non-negative, and the component function f_{n+1} is non-positive.

Because of (b), the component functions $f_1,\dots,f_n$ are interpreted as **commodity demand functions;** f_{n+1} is the **individual labour supply function,** and p_{n+1} is the **wage rate.** The variable m represents **non-earned income.** Obviously, the special case of pure consumption is included in the definition.

In Subsection 3.1 we will continue Hildenbrand's proof. It is straightforward to see that the matrix of per capita income effects is, in general, not positive semi-definite if the individual demand function f has property (b), irrespective of the density $\rho(m)$. However, when taking labour supply decisions into the model, it is much more interesting to integrate f(p,m) with respect to the wage rate. This will be done in Subsection 3.2.

Notice that the distribution of personal income is now endogenous. The function $(p_{n+1},m) \longmapsto m - p_{n+1} f_{n+1}(p,m)$ transforms a given distribution of wage rates and non-labour incomes into a distribution of total personal income. We will return to this point in Subsection 3.2.

In Section 2 we have tacitly assumed that we may reverse the order of integration and differentiation. It follows from the dominated convergence theorem (e.g., Loève, 1977, pp. 126-127) that one may differentiate under the integral sign if (i) the probability measure μ is concentrated on a finite interval [a,b] and (ii) the integrand f is a continuously differentiable function. We will therefore assume that the wage rate distribution has property (i).

3.1. Distribution of Non-Labour Income

By the weak axiom of revealed preference f has a negative semi-definite substitution matrix; for the relation between the weak axiom and the definiteness of the substitution matrix it is inessential whether or not f is positive-valued.

To complete Hildenbrand's proof, we have to show that the matrix, A, of per capita income effects is positive semi-definite if m is uniformly distributed over the interval [0,b]. The matrix $A=(a_{ij})$ is given by

$$a_{ij} = \frac{1}{b}\cdot\int_0^b f_j(p,m)\cdot\delta_m f_i(p,m)\,dm \qquad (i,j=1,\ldots,n+1).$$

Pick any vector $v\varepsilon R^{n+1}$. Calculating the quadratic form $v\cdot A\cdot v$ yields

$$v\cdot A\cdot v = \frac{1}{2b}\cdot\{(v\cdot f(p,b))^2 - (v\cdot f(p,0))^2\}.$$

If $f\geq 0$ (i.e., pure consumption), then the Walras identity implies that $f(p,0)=0$. Hence $vAv\geq 0$, i.e., A is positive semi-definite. However, if the behaviour described by f includes labour supply, then typically $f(p,0)\neq 0$, and hence we cannot conclude any more that A is positive semi-definite. Whether or not F is monotone will now depend on the matrix of the per capita substitution effects; since the substitution matrix is negative semi-definite, $\delta F(p)$ may be negative semi-definite even if A is not positive semi-definite.

The elements along the main diagonal of A are obtained by setting v = i-th unit vector of R^{n+1}. Thus,

$$a_{ii} = \frac{1}{b}\cdot\int_0^b f_i(p,m)\cdot\delta_m f_i(p,m)\,dm = \frac{1}{2b}\cdot\{f_i(p,b)^2 - f_i(p,0)^2\}.$$

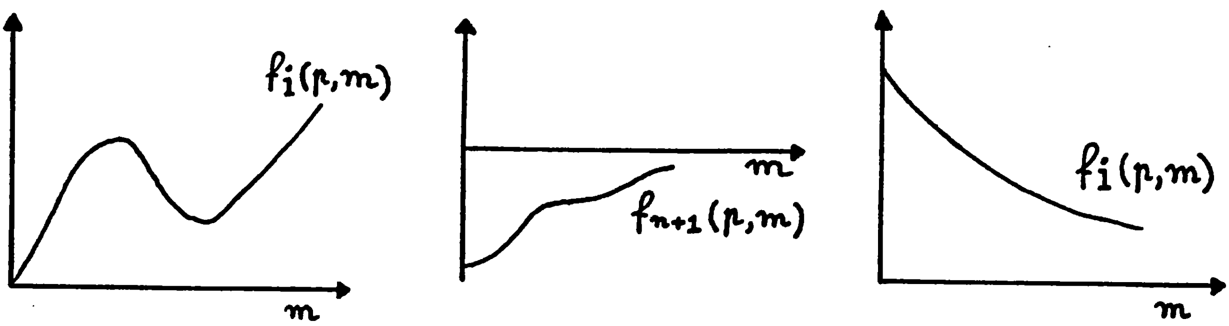
Non-negativity of f together with $f(p,0)=0$ implies that the sign of $\delta_m f_i(p,m)$ cannot be negative over the whole interval [0,b], i.e., the Engel curve $m \longmapsto f_i(p,m)$ is either monotone increasing on R_+ or first increasing and then decreasing (resp. "periodically" decreasing and increasing); see Figure 1(a). This behaviour of the Engel curve, in turn, implies that over any given interval [0,b] the average value a_{ii} of the individual income effect $f_i(p,m)\cdot\delta_m f_i(p,m)$ is non-negative. The individual labour supply function changes matters drastically. Now the mean income effect a_{ii} may have the "wrong" sign irrespective of the distribution of non-earned income, i.e., globally decreasing functions $f_i(p,\cdot)$ cannot be excluded any more. Indeed, the standard assumption that leisure is a *normal good* means that $\delta_m f_{n+1}(p,m)$ is positive for all $m\geq 0$ (recall the sign convention made at the beginning of the section). Hence, the integral

$$\int f_{n+1}(p,\cdot)\,\delta_m f_{n+1}(p,\cdot)\,d\mu$$

may be negative for any given distribution μ [see Figure 1(b)]. Since the household will have positive earned income at $m=0$, the demand for a consumer good may be a strictly decreasing function of m, i.e., for some $i\in\{1, \ldots, n\}$ the integral

$$\int f_i(p,\cdot)\,\delta_m f_i(p,\cdot)\,d\mu$$

may have a negative sign irrespective of μ [see Figure 1(c)].



(a) (b) (c)

Figure 1

Notice that $\delta_{p_i} F_i(p)$ (i=1,...,n+1) has a negative sign for any given decreasing density $\rho(m)$ if $f_i(p,m) \geq f_i(p,0)$ (i=1,...,n) and $f_{n+1}(p,m) \leq f_{n+1}(p,0)$ for all m (see note 3). Figure 2 illustrates this behaviour of the function f(p,·).

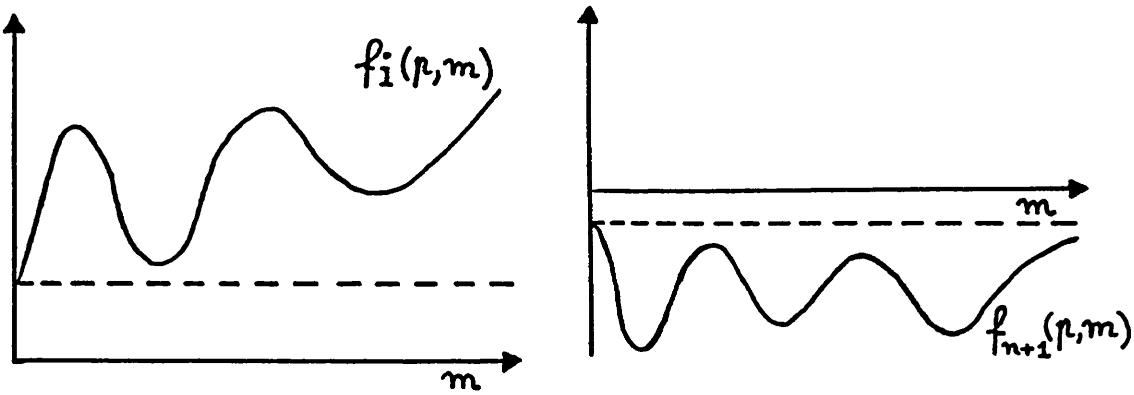


Figure 2

We will now prove Proposition 1 (p. 40). As already mentioned, the same proof can be found in Freixas and Mas-Colell (1987) who explore under which conditions on the individual demand function f (as defined in Section 2, i.e., f is non-negative) the aggregate demand function F will satisfy the weak axiom of revealed preference for any given income distribution.

Proof of Proposition 1: If the individual demand function f is derived from L-shaped preferences [i.e., the utility function underlying f is of the form $u(x_1,...,x_n) = \min\{u_1(x_1),...,u_n(x_n)\}$, where the functions u_i are strictly increasing], then its substitution matrix vanishes. This is the worst that can happen if a given demand function f satisfies the weak axiom. Suppose the density function ρ: $[0,b] \longrightarrow R_+$ is strictly increasing on $[m_1,m_2]$, where $0<m_1<m_2<b$. The question arises whether there exist L-shaped preferences so that the corresponding aggregate demand function F is not monotone. This is indeed the case. Consider without loss of generality the case n=2; pick any vector v in R^2 with a positive and a negative component and construct L-shaped preferences in such a manner that the individual demand function f has the following properties: (a) vf(p,m)=0 for all

$m \notin [m_1, m_2]$, and (b) $vf(p,m) \neq 0$ on a subinterval of $[m_1, m_2]$ (see Figure 3). Let $A = A(p)$ be the matrix of per capita income effects corresponding to f and ρ, i.e., the element a_{ij} of A is given by

$$a_{ij} = \int_0^b f_j(p,m) \cdot \delta_m f_i(p,m) \rho(m) dm.$$

Let $x \varepsilon R^n$. By partial integration, we obtain

$$xAx = \frac{1}{2} \cdot [(xf(p,m))^2]_0^b - \frac{1}{2} \cdot \int_0^b (xf(p,m))^2 \rho'(m) dm.$$

Because of (a), we have at x=v

$$xAx = -\frac{1}{2} \cdot \int_{m_1}^{m_2} (vf(p,m))^2 \rho'(m) dm.$$

Since ρ is strictly increasing on $[m_1, m_2]$ and f(p,m) satisfies (b), the integral on the right-hand side is positive. Thus, the matrix A(p) is not positive semi-definite; hence, the Jacobian matrix $\delta F(p)$ is not negative semi-definite. Q.E.D.

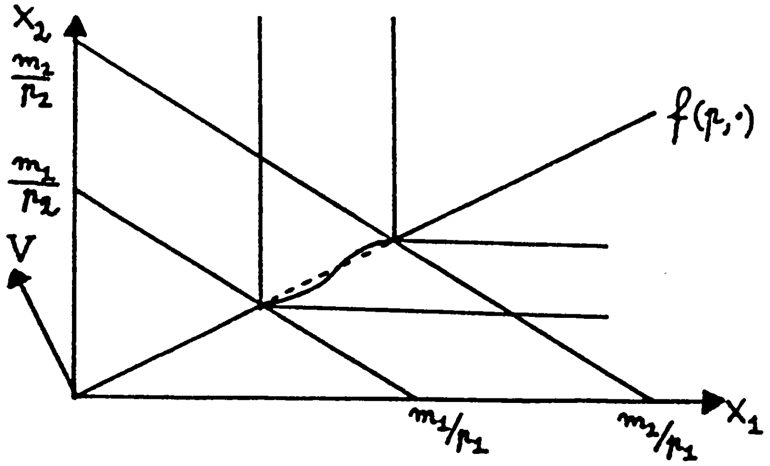


Figure 3

We now turn to the case where otherwise identical individuals face different wage rates; we will assume that all individuals receive the same non-labour income.

3.2. Distribution of Wage Rates

Let Φ: $P\times R_+ \longrightarrow R^{n+1}$ be an individual demand function and μ be a distribution of wage rates, i.e., μ is a probability measure on R_+ with compact support which can be represented by a density function ρ. To distinguish easily between commodity demand and labour supply, we define:

$w := p_{n+1}$, $p := (p_1,\ldots,p_n)$,

$f(p,w,m) := (\Phi_1(p,w,m),\ldots,\Phi_n(p,w,m))$ (commodity demand),

and $l(p,w,m) := -\Phi_{n+1}(p,w,m)$ (labour supply).

We are interested in two questions:

(1) Under what conditions is the per capita commodity demand function monotone in p?

(2) Under what conditions does a proportional or absolute rise in all wage rates lead to an increase in per capita labour supply?[5]

We first turn to the market demand function

$$F(p) = \int f(p,w,m)\,d\mu(w).$$

3.2.1. Per Capita Commodity Demand

Let $S\Phi(p,w,m)$ denote the substitution matrix of $\Phi=(f,-l)$ at (p,w,m). Since $\Phi=(f,-l)$ satisfies the the weak axiom, the matrix $S\Phi(p,w,m)$ is negative semi-definite for all (p,w,m).[6] The substitution matrix of the commodity demand function f with respect to p is obtained by deleting the last row and the last column from $S\Phi(p,w,m)$. Hence, the substitution matrix of the map $p \longmapsto f(p,w,m)$ is negative semi-definite for all p, w, and m. Since μ is concentrated on a finite interval, we may reverse the order of integration and differentiation. Thus, we can again decompose the Jacobian matrix of F into a negative semi-definite matrix of per capita substitution effects and a matrix of per capita income effects. Denoting the latter matrix by $A=(a_{ij})$, we have

$$a_{ij} = \int f_j(p,w,m)\cdot\delta_m f_i(p,w,m)\,d\mu(w) \quad \text{for all } i,j=1,\ldots,n.$$

We want to relate this matrix to the matrix of income effects explored by Hildenbrand (1983). To do this, we need the following crucial

Assumption 1: Consumption and labour are *(weakly) separable,* i.e., there exists a continuously differentiable homogeneous individual demand function f^s: $P\times R_+ \longrightarrow R^n_+$ (as defined in Section 2) so that for all (p,w,m)

$$f(p,w,m) = f^s(p,b(p,w,m)),$$

where $b(p,w,m)=m+wl(p,w,m)$; that is the map $(p,w,m)\longmapsto b(p,w,m)$ is the *individual income function.*

Separability between consumption and labour supply means that the wage rate affects demand behaviour only via its impact on individual income. If the individual demand function $\Phi=(f,-l)$ is derived from utility maximisation (what we have not assumed here), then the separability assumption implies that the underlying utility function U(x,l), where $x = (x_1,\ldots,x_n)$ represents the consumption goods, is of the form U(x,l)=u[v(x),l] (and hence the marginal rate of substitution between any two commodities is independent of the amount of labour supplied). Maximising U(x,l) subject to the budget constraint $px\leq wl+m$ (and an additional constraint on the maximum amount of labour the individual can supply) leads to the commodity demand function f(p,w,m) and the labour supply function l(p,w,m). The special form of U(x,l) allows now to decompose the overall maximisation problem into two sub-maximisation problems: Suppose the individual has already decided to supply l units of labour. Then his income is given by y=wl+m. The best the consumer can do now is to maximise v(x) subject to the constraint $px\leq y$. This produces the conditional demand function $f^s(p,y)$. Maximising U(x,l) with respect to l subject to the constraints $x=f^s(p,y)$ and y=wl+m, we obtain the optimal labour supply decision l(p,w,m). Plugging y=wl(p,w,m)+m into $f^s(p,y)$ yields the optimal consumption decision f(p,w,m); see Barten and Böhm (1981, pp. 392-394, 399-401) for a more formal discussion.

We will call the function $w \longmapsto b(p,w,m)$ **individual earnings function.** In the following we will frequently drop the variable m since non-labour income is fixed.

Let us first show how the distribution of personal income depends on the wage rate density ρ and the individual earnings function $b(p,\cdot)$. The following assumption guarantees that the distribution of personal income has a density which depends in a very simple manner on ρ and $b(p,\cdot)$.

Assumption 2: The function $b(p,\cdot,m)$ is strictly increasing on R_{++}. Furthermore, $\delta_w b(p,w,m)>0$ for all w in R_{++}; and $\lim_{w \to 0+} \delta_w b(p,w,m)>0$.

We have defined the labour supply function $l(p,w,m)$ only for strictly positive wage rates w. We define $b(p,w)$ at $w=0$ by $b(p,0)=\lim_{w \to 0+} b(p,w)$; $\delta_w b(p,0)$ is defined by $\delta_w b(p,0)=\lim_{w \to 0+} \delta_w b(p,w)$. Notice, if a strictly increasing function $g: R_+ \longrightarrow R_+$ does not satisfy $g'(x)>0$ for all x in R_+, one only has slightly to change g in order to obtain $g'(x)>0$ on R_+; the same applies if $g: R_+ \longrightarrow R_+$ is identically zero on $[0,b]$ $(b>0)$ and strictly increasing on $[b,\infty[$.

The income distribution is given by the image of the distribution of w under the map $w \longmapsto b(p,w)$, i.e., by the probability distribution of the "random variable" $b(p,\cdot)$ (see, e.g., Loève, 1977, p. 168). We denote the distribution of $b(p,\cdot)$ by $\mu_{b(p,\cdot)}$; $b^{-1}(p,\cdot)$ denotes the inverse of $b(p,\cdot)$.

Proposition 2: *Let* ρ *be the density of* μ. *Suppose* $b(p,\cdot)$ *satisfies assumption 2.* Then $\mu_{b(p,\cdot)}$ *has a density function* $\tilde{\rho}(p,\cdot)$: $R_+ \longrightarrow R_+$ *which is given by*

$$\tilde{\rho}(p,y) = \rho(b^{-1}(p,y)) \cdot \frac{1}{\delta_w b(p,b^{-1}(p,y))}$$

for all y *in* $b(p,R_+)$, *and* $\tilde{\rho}(p,y) = 0$ *otherwise.*

Remark: Clearly, $b(p,R_+)=[b(p,0),\infty[$ if the earnings function $b(p,\cdot)$ is not bounded. Since we have assumed that the wage rate distribution is concentrated on a finite interval $[w_1, w_2]$, the income distribution is concentrated on the interval $[y_1, y_2]$, where $y_i = b(p, w_i)$, $i=1,2$.

Proof of Proposition 2: We have to show that for all $x \geq 0$

$$\mu_{b(p,\cdot)}([0,x]) = \int_0^x \tilde{\rho}(p,y)\,dy.$$

By definition of $\mu_{b(p,\cdot)}$, we have for all x in $b(p, R_+)$

$$\mu_{b(p,\cdot)}([0,x]) =$$

$$= \mu(\{w \in R_+ : b(p,w) \leq x\})$$

$$= \mu(\{w \in R_+ : w \leq b^{-1}(p,x)\}) \qquad \text{(assumption 2)}$$

$$= \int_0^{b^{-1}(p,x)} \rho(w)\,dw \qquad \text{(μ has the density ρ)}$$

$$= \int_{b(p,0)}^x \rho(b^{-1}(p,y)) \cdot \frac{1}{\delta_w b(p, b^{-1}(p,y))}\,dy,$$

where the last equality is obtained by making the substitution $w = b^{-1}(p,y)$. If x is not in $b(p, R_+)$, $\mu_{b(p,\cdot)}([0,x])$ is either equal to 1 or equal to 0. By definition of $\tilde{\rho}(p,\cdot)$, the last integral is equal to

$$\int_0^x \tilde{\rho}(p,y)\,dy.$$

Thus, $\tilde{\rho}(p,\cdot)$ is the density of the income distribution. Q.E.D.

The density $\tilde{\rho}(p,\cdot)$ may be interpreted Hildenbrand's density of individual incomes if all members of the population receive the same non-labour income m. We will come back to Proposition 2 at the end of the section.

We now relate the matrix A to the matrix of per capita income effects explored by Hildenbrand. To do this, we need the marginal dependence of total income on non-labour income, i.e., the partial derivative $\delta_m b(p,w,m)$. For fixed p and m, we define a function g from R_+ into R by

$$g(y) = \begin{cases} \delta_m b(p,b^{-1}(p,y),m), & \text{if } y \text{ in } b(p,R_+) \\ 0, & \text{otherwise.} \end{cases}$$

Thus, if the individual described by $\Phi=(f,-1)$ has total income y, g(y) is the rate of change in his total income resulting from a small increase in non-labour income m; if m=0, then $\delta_m b(p,w,m)$ is defined as $\delta_m b(p,w,0) = \lim_{m \to 0+} \delta_m b(p,w,m)$. The following proposition expresses the matrix A in terms of the probability measure $\mu_{b(p,\cdot)}$ and the functions f^s and g.

Proposition 3: *Let* μ *be a distribution of wage rates and* $\Phi=(f,-1)$ *be an individual demand function satisfying assumption 1. Then*

$$\int f(p,w,m)d\mu(w) = \int f^s(p,\cdot)d\mu_{b(p,\cdot)}.$$

Suppose the earnings function satisfies assumption 2. Let $\bar{\rho}(p,\cdot)$ *be the density of* $\mu_{b(p,\cdot)}$, *and let* $g(\cdot)$ *be as defined above. Then the matrix* $A = (a_{ij})$ *can be written as*

$$a_{ij} = \bar{g}\cdot\int f_j^s(p,y)\cdot\delta_y f_i^s(p,y)\cdot\bar{\rho}(p,y)dy$$

$$+ \operatorname{cov}(f_j^s(p,\cdot)\delta_y f_i^s(p,\cdot),g(\cdot))$$

where the second summand denotes the covariance of $f^s_j(p,\cdot)\delta_y f^s_i(p,\cdot)$ *and* $g(\cdot)$ *with respect to the measure* $\mu_{b(p,\cdot)}$; *and* $\bar{g} = \int g(y)\bar{\rho}(p,y)dy$, *i.e.,* $\bar{g}$ *is the mean value of* $g(\cdot)$ *with respect to* $\mu_{b(p,\cdot)}$.

Proof: The first part of the proposition follows immediately from the definition of $\mu_{b(p,\cdot)}$ and is a standard result in the theory of integration (see, e.g., Loève, 1977, p. 168). Turning to the second part, we have:

$$a_{ij} = \int f_j^s(p,b(p,w,m))\cdot\delta_m f_i^s(p,b(p,w,m))d\mu(w) \quad \text{(assumption 1)}$$

$$= \int f_j^s(p,b(p,w,m))\cdot\delta_y f_i^s(p,b(p,w,m))\cdot\delta_m b(p,w,m)\,d\mu(w)$$

$$= \int f_j^s(p,y)\cdot\delta_y f_i^s(p,y)\cdot g(y)\,d\mu_{b(p,\cdot)} \quad \text{(def. of } g(\cdot) \text{ and } \mu_{b(p,\cdot)})$$

$$= \int f_j^s(p,y)\cdot\delta_y f_i^s(p,y)\cdot g(y)\cdot\bar{\rho}(p,y)\,dy \quad \text{(def. of density).}$$

By definition of the covariance of two random variables X and Y with respect to a probability measure P,

$$\mathrm{cov}(X,Y) = \int XYdP - \int XdP\cdot\int YdP.$$

Hence, letting $X = f^s_j(p,\cdot)\delta_y f^s_i(p,\cdot)$, $Y = g(\cdot)$ and $P = \mu_{b(p,\cdot)}$ completes the proof.

Example: Suppose an individual behaves as if maximising a function $U(l) = u(wl+m) - v(l)$ (which may depend on p), where u is concave and v is convex; $u'>0$ and $v'>0$. Then the labour supply function $l(w,m)$ is implicitly given by the first-order condition $u'(wl+m)w = v'(l)$. Consider the following two special cases: if $v(l) = \alpha l$, then $l(w,m) = u'^{-1}(\alpha/w)/w - m/w$ and hence $\delta_m b(p,w,m) \equiv 0$; if $u(y) = \alpha y$, then $l(w) = v'^{-1}(\alpha w)$ and therefore $\delta_m b(p,w,m) \equiv 1$.

Since the elements of the matrix $A=(a_{ij})$ may be written in the form

$$a_{ij} = \int f_j^s(p,y)\cdot\delta_y f_i^s(p,y)\cdot g(y)\cdot\bar{\rho}(p,y)\,dy,$$

the next proposition is an immediate consequence of Hildenbrand (1983, Theorem 3). Notice that $g(\cdot)$ may be identically zero and that $\bar{\rho}(p,\cdot)$ can only be decreasing on R_+ if m=0 since $b(p,w,m)\geq m$.

Poposition 4: *Let* $\Phi=(f,-l)$ *be an individual demand function which satisfies assumptions 1 and 2. Let* $\bar{\rho}(p,\cdot)$: $R_+ \longrightarrow R_+$ *be the density of* $\mu_{b(p,\cdot)}$. *Suppose* $g(\cdot)\bar{\rho}(p,\cdot)$ *is decreasing on* R_+ *and* $g(\cdot)\geq 0$. *Then*

(i) *the matrix* A *is positive semi-definite at* p;

(ii) *the Jacobian matrix of* F *is negative semi-definite at* p.

If the income density is monotone decreasing but depends on the price system p, Hildenbrand cannot conclude that the Jacobian matrix $\delta F(p)$ is negative semi-definite. The relation between Hildenbrand's approach and our approach is as follows: Let $\Phi=(f,-l)$ and f^s be individual demand functions (as defined in Section 2 and 3) such that $f(p,w,m) = f^s(p,b(p,w,m))$. Let $\rho(p,\cdot)$ be the density of the wage rate distribution at the prevailing price system p. Finally, let $\tilde{\rho}(p,\cdot)$ be the income density generated by the labour supply function $l(p,w,m)$ and the wage rate density $\rho(p,\cdot)$.

The two primitive concepts of Hildenbrand's approach are the individual demand function f^s and the income density $\tilde{\rho}(p,\cdot)$. Market demand is defined by $F(p)=\int f^s(p,y)\tilde{\rho}(p,y)dy$. Differentiating under the integral sign and applying the Slutzky decomposition, $\delta F(p)$ may be written as

$$\delta F(p) = \tilde{S} - \tilde{A} + \tilde{B}, \text{ where}$$

$$\tilde{s}_{ij} = \int \delta_{q_j} s_i(q,y)_{|q=p} \cdot \tilde{\rho}(p,y)dy,$$

$$\tilde{a}_{ij} = \int f^s_j(p,y) \cdot \delta_y f^s_i(p,y) \cdot \tilde{\rho}(p,y)dy$$

and

$$\tilde{b}_{ij} = \int f^s_i(p,y) \cdot \delta_{p_j} \tilde{\rho}(p,y)dy.$$

For given p and y, the function $q \longmapsto s(q,y)$ is the Slutzky compensated demand function corresponding to $f^s(p,y)$, i.e., $s(q,y)=f^s(q,qf^s(p,y))$. Here we have assumed that the (continuously differentiable) income density $\tilde{\rho}(p,\cdot)$ is concentrated on a finite interval $[y_1,y_2]$ which does not depend on p. If the range of $\tilde{\rho}(p,\cdot)$ depends on p, i.e., $[y_1,y_2] = [\alpha(p),\beta(p)]$, we obtain $\delta F(p) = \tilde{S} - \tilde{A} + \tilde{B} + C$, where the matrix $C=(c_{ij})$ is given by

$$c_{ij} = z_i(p,\beta(p)) \cdot \delta_{p_j} \beta(p) - z_i(p,\alpha(p)) \cdot \delta_{p_j} \alpha(p),$$

and $z(p,y) = f^s(p,y)\tilde{\rho}(p,y)$; see Courant and John (1974, pp. 76-77).

In general, the matrix $\tilde{B}$ may have any structure. To see this, notice first that the partial derivative of $\tilde{\rho}(p,y)$ with respect p_j cannot be negative (resp. positive) on the entire interval $[y_1,y_2]$ since $\tilde{\rho}(p,\cdot)$

integrates to 1 for any given p. Suppose the Engel curves $f^s_i(p,\cdot)$ (i=1, ...,n) are strictly increasing and that $\delta_{p_j}\tilde{\rho}(p,y)$ is non-positive on $[y_1,z]$ and non-negative on $[z,y_2]$, where z=z(j), j=1,...,n (the latter assumption is without loss of generality); see Figure 4 below. Then the elements $\tilde{b}_{ij}$ of $\tilde{B}$ are positive:

$$\tilde{b}_{ij} = \int f^s_i(p,y)\delta_{p_j}\tilde{\rho}(p,y)dy$$

$$> f^s_i(p,z)\cdot\int_{y_1}^{z}\delta_{p_j}\tilde{\rho}(p,y)dy + f^s_i(p,z)\cdot\int_{z}^{y_2}\delta_{p_j}\tilde{\rho}(p,y)dy$$

$$= 0.$$

Clearly, this implies that the matrix $\tilde{B}$ is not negative semi-definite. Hence, even if $\tilde{A}$ is positive semi-definite, $\delta F(p)$ may not be negative semi-definite.

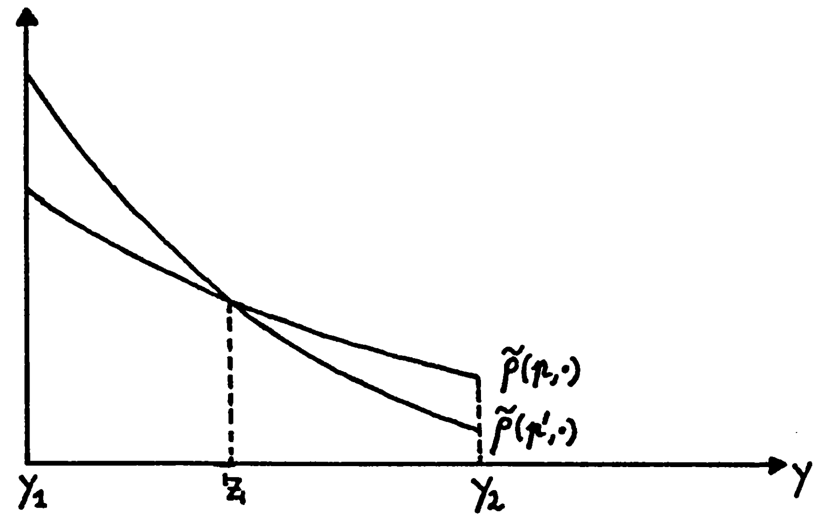


Figure 4

The primitive concepts of the present approach are the wage rate density $\rho(p,\cdot)$ and the function $\Phi=(f,-1)$; that is, $F(p)=\int f(p,w,m)\rho(p,w)dw$. We obtain

$\delta F(p) = S - A + B$, where

$$s_{ij} = \int \delta_{q_j} s_i(q,w,m)|_{q=p} \cdot \rho(p,w)\,dw,$$

$$a_{ij} = \int f_j^s(p,y) \cdot \delta_y f_i^s(p,y) \cdot g(y) \cdot \tilde{\rho}(p,y)\,dy \qquad \text{(Proposition 3)}$$

and

$$b_{ij} = \int f_i(p,w,m) \cdot \delta_{p_j} \rho(p,w)\,dw.$$

For fixed p, w and m, the function $q \mapsto s(q,w,m)$ is the Slutzky compensated demand function corresponding to $f(p,w,m)$, i.e., $s(q,w,m)=f(q,w,m')$, where $m'=qf(p,w,m)-wl(p,w,m)$. The income density $\tilde{\rho}(p,\cdot)$ depends on p via the dependence of individual labour supply on p (even if the wage rate density does not depend on p). In our approach, however, it is only the relationship between p and $\rho(p,\cdot)$ which may cause problems (i.e., the matrix B may "wipe out" structural properties of A). It is straightforward to verify that the matrix $\tilde{B}$ is related to the matrix B as follows:

$$\tilde{b}_{ij} = b_{ij} + \int [\delta_y f_i^s(p,b(p,w,m)) \cdot \delta_{p_j} b(p,w,m)] \cdot \rho(p,w)\,dw.$$

It is natural to ask under what conditions on the wage rate density ρ and the earnings function $b(p,\cdot)$ the income density $\tilde{\rho}(p,\cdot)$ is decreasing on an interval $[y_1,y_2]$. By Proposition 1,

$$\tilde{\rho}(p,y) = \rho(b^{-1}(p,y)) \cdot \delta_y b^{-1}(p,y).$$

Suppose ρ is uniform on the interval $[w_1,w_2]$. Let $R=w_2-w_1$. Then

$$\tilde{\rho}(p,y) = \frac{1}{R} \cdot \frac{1}{\delta_w b(p,b^{-1}(p,y))}$$

on the interval $[y_1,y_2]$, where $y_i=b(p,w_i)$, $i=1,2$; and $\tilde{\rho}(p,y)=0$ otherwise. Hence $\tilde{\rho}(p,\cdot)$ is (strictly) decreasing on $[y_1,y_2]$ if and only if $b(p,\cdot)$ is (strictly) convex on $[w_1,w_2]$; $b(p,\cdot)$ is (strictly) convex on $[w_1,w_2]$ if and only if the labour supply function $l(p,\cdot,m)$ is (strictly) increasing on $[w_1,w_2]$.

Postulating that the individual labour supply function is monotone increasing in w would be an extremely strong assumption. However, if ρ is sufficiently decreasing then $\bar{\rho}(p,\cdot)$ will be decreasing even if $l(p,\cdot,m)$ is strictly decreasing. One immediately verifies that the derivative $\delta_y\bar{\rho}(p,y)$ is non-positive if and only if

$$\frac{\rho'(a(y))}{\rho(a(y))} \leq - \frac{a''(y)}{a'(y)^2},$$

where $a(y)=b^{-1}(p,y)$.

Example: The *Pareto distribution* with parameters $w^*>0$ and $\beta>0$ is defined by the density function

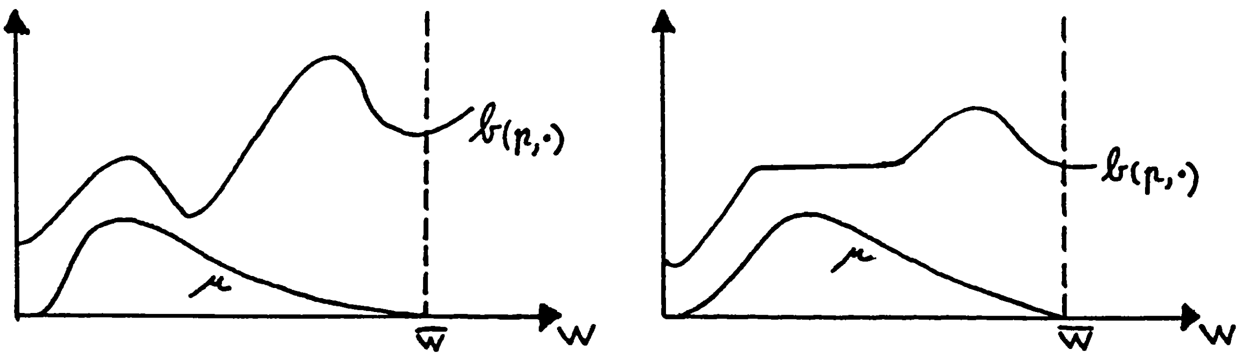
$$\rho(w) = \alpha\beta w^{-(1+\beta)}, \qquad w \geq w^*,$$

and $\rho(w)=0$ otherwise; $\alpha=(w^*)^\beta$ in order to ensure that ρ integrates to one.[7] If the wage rates are distributed according to the Pareto distribution, then the income density is decreasing on $[y_1,y_2]$, where $y_i=b(p,w_i)$ $(i=1,2)$ and $w^* \leq w_1 < w_2$, if and only if $a(y)=b^{-1}(p,y)$ satisfies

$$\frac{a(y)}{a'(y)} \cdot \frac{a''(y)}{a'(y)} \leq 1+\beta \qquad \text{for all } y \text{ in } [y_1,y_2].$$

The first quotient on the left-hand side of the inequality is positive for all y in $[y_1,y_2]$; the second quotient is positive at all points y where the earnings function is locally strictly concave, and smaller than or equal to zero otherwise.

Let us take stock. Obviously, assumption 1 implicitly underlies Hildenbrand (1983). We do not think that assumption 2 is restrictive.[8] For instance, individual income is always an increasing function of the wage rate if the demand function $\Phi=(f,-l)$ is derived from a quasi-concave utility function which admits an additively separable representation $u(x_1,\ldots,x_n,l) = u_1(x_1,\ldots,x_n) + u_2(l)$, where x_i stands for the i-th commodity. However, assumption 2 was only made for the purpose of exposition. If one merely wants to ensure that the income distribution has a density, then assumption 2 is not needed: By the Radon-Nikodym theorem (see, e.g., Loève, 1977, p. 133), $\mu_{b(p,\cdot)}$ has a μ-density if (and only if) any μ-null set is also a $\mu_{b(p,\cdot)}$-null set, i.e., $\mu_{b(p,\cdot)}$ is μ-continuous. Since we have assumed that μ can be represented by a Lebesgue density, μ-continuity of $\mu_{b(p,\cdot)}$ implies that $\mu_{b(p,\cdot)}$ has also a Lebesgue density. The distribution $\mu_{b(p,\cdot)}$, in turn, is μ-continuous, if the function $b(p,\cdot)$ is not locally constant on the support of μ (see Figure 5).



$b(p,\cdot)$ is μ-continuous

$b(p,\cdot)$ is not μ-continuous

Figure 5

We now turn to the response of per capita labour supply to an increase in the wage level; in the following we will also drop the variable p.

3.2.2. Per Capita Labour Supply

If leisure is a desired commodity, the individual will supply no labour at $w=0$; if the wage rate is positive, however small, the individual may want to supply at least the amount $\bar{l}>0$. Thus, the function $l(w)$, when defined on the entire R_+, may be discontinuous at $w=0$ (see Figure 6). We define $l(w)$ at $w=0$ by $l(0)=\lim_{w\to 0+} l(w)$.

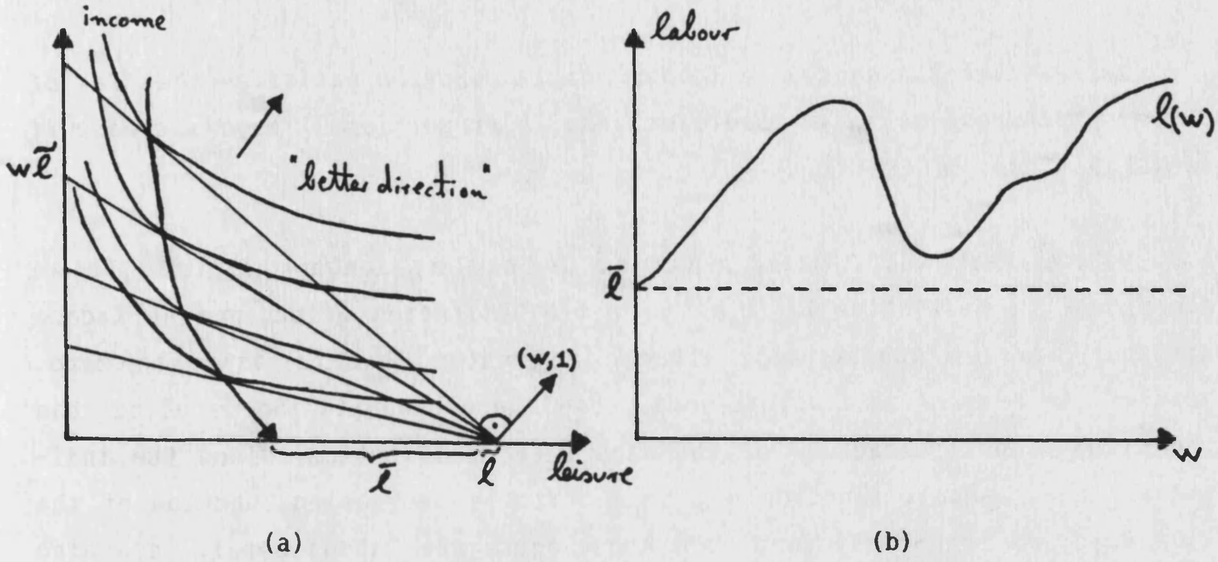


Figure 6

Given the wage rate distribution μ and the individual labour supply function l(w), per capita labour supply, L, is defined by

$$L = \int l(w)d\mu(w).$$

If we add to all wage rates the amount a>0, per capita labour supply becomes

$$L(a) = \int l(w+a)d\mu(w) \qquad (a \in R_+).$$

If we increase all wage rates by (α-1)·100%, we have to substitute αw for w+a. Hence, the rate of change of L resulting from a small absolute increase in all wage rates is given by

$$\delta_a L(a)_{|a=0} = \int l'(w)d\mu(w),$$

and the rate of change of L resulting from a small proportional increase is given by

$$\delta_\alpha L(\alpha)_{|\alpha=1} = \int l'(w)wd\mu(w).$$

We say that the aggregate labour supply function satisfies the "law of supply" with respect to an absolute (resp. a proportional) wage increase if $\delta_a L(0) \geq 0$ *[resp.* $\delta_\alpha L(1) \geq 0$*].*

Notice that the Slutzky equation is now no longer helpful. If we decompose the derivative $\delta_w l(w,m)$ into a substitution effect and an income effect, then the substitution effect is greater than or equal to zero. However, if leisure is a normal good, then the per capita income effect has the wrong sign irrespective of the wage rate distribution. Since the individual labour supply function may be a strictly decreasing function of the wage rate, we cannot say very much about aggregate labour supply (see also Subsection 3.1). The following observation is an immediate consequence of the *second mean value theorem.*

Proposition 5: *Let* $z \geq 0$. *Let* D_z *denote the class of density functions* $g: R_+ \longrightarrow R_+$ *which have the following properties:* (i) $g(w)=0$ *for all* x *in* $[0,z[$; (ii) g *is decreasing on* $[z,\infty[$; *and* (iii) g *is concentrated on a finite interval. Then the following holds:*

(a) $\delta_a L(0) \geq 0$ *for all* g *in* D_z *if and only if*

$$\text{(A)} \quad l(w) \geq l(z) \qquad \textit{for all } w \geq z;$$

(b) $\delta_\alpha L(1) \geq 0$ *for all* g *in* D_z *if and only if*

$$\text{(P)} \quad l(w)w - l(z)z \geq \int_z^w l(w)\,dw \qquad \textit{for all } w \geq z.$$

Thus, if $z=0$, *then* $\delta_\alpha L(1) \geq 0$ *for all* g *in* D_z *if and only if*

$$\text{(P')} \quad l(w) \geq \frac{1}{w} \cdot \int_0^w l(w)\,dw \qquad \textit{for all } w > 0.$$

Proof: Let $g: [z,\infty[\longrightarrow R_+$ be decreasing and concentrated on $[z,b]$. By the second mean value theorem (see note 3) there exist x, y in $[z,b]$ such that

$$\delta_a L(0) = g(z) \cdot \int_z^x l'(w)\,dw$$

and

$$\delta_\alpha L(1) = g(z)\cdot\int_z^y l'(w)wdw.$$

Hence,

$$\delta_a L(0) = g(z)\cdot[l(x) - l(z)]$$

and, by partial integration,

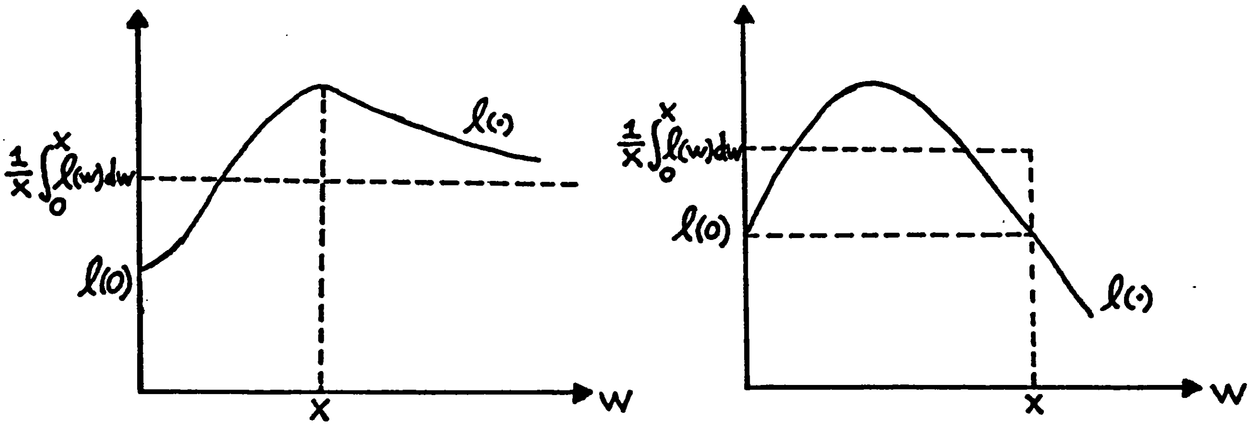
$$\delta_\alpha L(1) = g(z)\cdot[l(y)y - l(z)z - \int_z^y l(w)dw].$$

If g is the density of the uniform distribution on [z,b], then x=y=b. This completes the proof.

Thus, if the wage rate density is decreasing, our conclusion is essentially that the "law of supply" holds if individuals working for firms which pay wage rates $w>z$ do not work less than those located at the bottom of the wage rate distribution (see Figure 6b above). In view of the "efficiency wage" literature (see, e.g., the articles collected in Akerlof and Yellen, 1986) such a hypothesis is perhaps not unplausible. We will see in Chapter 3 whether or not empirical labour supply curves satisfy (A).

Notice that the right-hand side of (P') is the average value of the function $l(\cdot)$ in the interval [0,w]. Hence, (P') says that the labour supply of an individual earning the wage rate w must not be smaller than the average labour supply (with respect to the uniform distribution on [0,w]) of those individuals receiving wage rates between 0 and w (see Figure 7a on the next page). Property (A) does not imply (P) (consider, e.g., the labour supply function $l(w)=1-(1-w)^2$ and a uniform wage rate distribution on $[0,2-\tau]$, $\tau>0$); but (P') implies (A) (see Figure 7b).

Let $z>0$. Let D'_z denote the class of density functions $g: R_+ \longrightarrow R_+$ which have the following properties: (i) $g \in D_z$ and (ii) $wg(w)$ is decreasing on the interval $[z,\infty[$ (e.g., the densities of the on [z,b] (b>0) truncated Pareto distributions belong to D'_z).[9)] Applying the second mean value theorem to wg(w), one sees that of $\delta_\alpha L(1)$ has a positive sign for all g in D'_z if and only if (A) is satisfied.



(a) (b)

Figure 7

We close the section with an example that shows that the market demand function may be monotone decreasing in the commodity prices while we cannot say very much about the response of per capita labour supply to an increase in the wage level:
Consider an additively separable utility function, u, with constant marginal disutility of effort, $\beta>0$, i.e.,

$$u(x_1,\ldots,x_n,l) = v_1(x_1)+\ldots+v_n(x_n) - \beta l,$$

where the functions v_i are increasing and strictly concave; x_i stands for commodity i, and l denotes labour.

The "rational" individual described by u(x,l) maximises his utility subject to the budget constraint px=wl+m. From the first order conditions one obtains

$$f_i(p,w,m)=(v_i')^{-1}(\beta\cdot\frac{p_i}{w}), \quad i=1,\ldots,n.$$

Thus the market demand function is monotone decreasing irrespective of the wage rate distribution. (That additive utility functions with decreasing marginal utilities lead to market demand functions satisfying the "law of demand" is, of course, well-known and has already been recognised by

Walras, 1874) However, the behaviour of the individual labour supply function

$$l(p,w,m) = \frac{pf(p,w)}{w} - \frac{m}{w}$$

depends on the curvature of the functions v_i $(i=1,\ldots,n)$. It is not difficult to verify the following: Let $m=0$, $n=1$ and $v=v_1$. Then $\delta_w l(p,w)>0$ if and only if the function $z(x):=(v')^{-1}(x)$ satisfies $z(x)+x\cdot z'(x)<0$ at $x=\beta p/w$. For instance, if $v(x)=\sqrt{x}$ then this condition is satisfied for all $x\geq 0$; if $v(x)=-1/x$, then $z(x)+x\cdot z'(x)>0$ for all $x>0$, and hence $l(p,w)$ will be a strictly decreasing function of the wage rate.

4. Final Remarks

If one interprets $f(p,w,m)$ as the mean demand at the prevailing price system p of all individuals (of a given population) receiving the wage rate w and non-labour income m, then the mean income effect matrix $A=(a_{ij})$, where

$$a_{ij} = \iint f_j(p,w,m)\cdot\delta_m f_i(p,w,m)\cdot\rho(w,m)\,dw\,dm \quad (i,j=1,\ldots,n),$$

can be estimated from a sample of a consumption sector. To estimate A, one has to estimate the joint density of w and m, and one has to regress the expenditure on any commodity i on w and m. The matrix $\tilde{A}=(\tilde{a}_{ij})$, where

$$\tilde{a}_{ij} = \int f^s_j(p,y)\cdot\delta_y f^s_i(p,y)\cdot\tilde{\rho}(y)\,dy \quad (i,j=1,\ldots,n),$$

was analysed by K. Hildenbrand and W. Hildenbrand (1986) for an aggregated demand system of 11 commodities. It turned out that the estimated matrix is '"approximately" positive semi-definite' (p. 267). Hildenbrand (1989a) reports new results and concludes: "If the law of aerodynamics were founded as I have here founded the Law of Demand would I then take a plane?...

Perhaps I would feel somewhat uncomfortable, yet, I think, I would take a chance" (pp. 275-276).

Nevertheless, it would be interesting to estimate the matrix A and to compute its eigenvalues (if all eigenvalues are greater than zero, the estimated matrix is positive definite). Firstly, we have seen in Section 3 that the price dependence of the income distribution may cause more problems than the price dependence of the wage rate distribution. Secondly, there is no clear empirical evidence that consumption and labour are separable; see the time-series study by Abbott and Ashenfelter (1976, 1979), the pooled cross-section study by Browning and Meghir (1989) and the cross-section studies by Atkinson and Stern (1981), Blundell and Walker (1982, 1986) and Kaiser (1990).

One may try to characterise the class of joint distributions of preference relations, wage rates and non-labour incomes which lead to aggregate commodity demand and labour supply functions satisfying the "law of supply and demand". In the following two chapters we will be less ambitious. In Chapter 2 we will estimate the wage rate density ρ, and in Chapter 3 we will estimate the labour supply function $l(w)$. We will then compute $\delta_\alpha L(0)$ and $\delta_\alpha L(1)$.

5. Notes

1) It appears that the hypothesis is very well supported by empirical evidence; see, e.g., the nice discussion of inter-industry wage differentials in Thaler (1989) and the references given there.

2) The usual and somewhat stronger formulation of the weak axiom is: if $f(p,m)\neq f(p',m')$ and $pf(p',m')\leq m$, then $p'f(p,m)>m'$. The axiom was first used by Wald (1935, 1936a, 1936b) who proved the existence of market equilibria under the assumption that the market demand function satisfies the weak axiom of revealed preference. Later Samuelson (1938) used the axiom as a new foundation for the theory of individual consumption behaviour. Houthakker (1950) established the relation between the weak axiom and the utility hypothesis. The standard reference for the "demand theory of the weak axiom" is Kihlstrom et. al. (1976); W. Hildenbrand and M. Jerison (1988) simplify proofs. The volume of readings Chipman et. al. (1971) contains extensive surveys of demand theory.

3) The "second mean value theorem" states that for any continuous real-valued function f and any decreasing (resp. increasing) function g, both defined on an interval [a,b], there exists a point x_0 in [a,b] such that

$$(*) \qquad \int_a^b f(x)g(x)dx = g(a)\cdot\int_a^{x_0} f(x)dx + g(b)\cdot\int_{x_0}^b f(x)dx.$$

A proof can be found, e.g., in Apostol (1974, Theorem 7.37, p. 165). Notice, if g is decreasing on [a,b] with $g(x)\geq 0$, one can replace g by the function g^* defined by $g^*(x)=g(x)$ if $a\leq x<b$ and $g^*(b)=0$. Substituting g^* for g does not change the value of the integral on the left-hand side of (*).

4) Of course, this strong assumption on the distribution of personal income is only needed as long as one does not want to restrict the class of permissible individual demand functions. Imposing no restrictions on individual demand functions is an extreme starting point. The other extreme would be to restrict attention to homothetic preferences, i.e., to demand functions $f(p,m)$ which can be written as $f(p,m)=g(p)\cdot m$. Recall that we have shown in Chapter 0 (Proposition 5, p. 27) that the weak axiom and the budget identity imply that the function g is monotone (it is not difficult to verify that already the weak version of the weak axiom implies the monotonicity of g). The reader may find the following articles interesting:
The class of preference relations leading to demand functions $f(p,m)$ which are monotone in p for any given m was characterised by Kannai (1987); see also Mitjuschin and Polterovich (1978). Formalising Marshall's old idea that the income effects $f_i(p,m)\cdot\delta_m f_i(p,m)$ $(i=1,\ldots,n)$ will be small if the proportion of total income m spend on any commodity i is small, Vives (1987) states conditions which imply that the partial demand functions $f_i(p,m)$ are decreasing in their own price. Grandmont (1987) assumes that people have the same income but that they differ with respect to their preferences. He then places restrictions on the "shape" of the distribution of preference relations which lead to a monotone market demand function; see also Grandmont (1992). Freixas and Mas-Colell (1987) study under what conditions on the function $f(p,m)$ the market demand function $F(p)$ satisfies the weak axiom of revealed preference (i.e., if $F(p) \neq F(p')$ and $pF(p') \leq$ mean income, then $p'F(p) >$ mean income) irrespective of the shape of the income density. Chiappori (1985) considers functions $f(p,m)$ which can be written in the form

$$f(p,m) = \sum g_i(p)\tau_i(m),$$

where the g_i are vector-valued functions, and the τ_i are real-valued functions. He then states sufficient conditions for the existence of a decreasing function $\rho^*: R_+ \longrightarrow R_+$ which solves the integral equation

$$\int f(p,\cdot)d\mu = \int f(p,m)\,\rho^*(m)dm,$$

where μ is a given income distribution. Chiappori does not, however, explore whether there exist individual demand functions (as defined in Section 2) which are non-linear in m and which can be decomposed in the above way (it should be mentioned that his examples are not consistent with the definition of an individual demand function in Hildenbrand, 1983). In applied demand theory it is frequently assumed that the market demand function may be expressed as a function of mean income; a discussion can be found in Hildenbrand (1985a).

5) In textbooks on macroeconomics it is usually assumed that aggregate labour supply is an increasing function of the wage level. For instance, Blanchard and Fischer (1989, p. 518) specify the following labour supply function in order to study supply shocks: $n^s=\delta(w-p)$, where $\delta\geq 0$. Here n^s, w and p are the logarithms of labour supply, the nominal wage and the price level, respectively. The labour supply function is part of

a model (i.e., a system of five equations) that as the authors write "has played a central role in the analysis of economic fluctuations..." (p. 518).

6) In the literature it is usually assumed that $m>0$, i.e., one proves the following: if Φ is a demand function which satisfies the weak axiom of revealed preference, then the substitution matrix is negative semi-definite for all $p>>0$ (here p includes the wage rate w) and $m>0$. In Subsection 3.2.1 it is natural to set $m=0$. Of course, the substitution matrix $S\Phi(p,m) = (s_{ij})$ is also negative semi-definite at $(p,m)=(p,0)$: The elements s_{ij} are defined by

$$s_{ij} = \delta_{p_j}\Phi_i(p,m) + \Phi_j(p,m)\delta_m\Phi_i(p,m).$$

At the point $(p,0)$, $\delta_m\Phi_i(p,m)$ is defined as $\delta_m\Phi_i(p,0) = \lim_{m\to 0+}\delta_m\Phi_i(p,m)$, and $\delta_{p_j}\Phi_i(p,0)$ is defined as $\delta_{p_j}\Phi_i(p,0) = \lim_{m\to 0+}\delta_{p_j}\Phi_i(p,m)$. Hence, $s_{ij}(p,m)\longrightarrow s_{ij}(p,0)$ as m tends to zero, and therefore $x\cdot S\Phi(p,m)\cdot x \longrightarrow x\cdot S\Phi(p,0)\cdot x$, as $m\longrightarrow 0$, for all $x\in R^{n+1}$. Thus, $x\cdot S\Phi(p,0)\cdot x \le 0$ since $x\cdot S\Phi(p,m)\cdot x \le 0$ for all $m>0$. Hence, $S\Phi(p,0)$ is negative semi-definite.

7) For instance, a "large" firm which is pyramidally organised, with a constant "span of control" $s>1$ and a constant wage differential of $(d-1)\cdot 100\%$ $(d>1)$ between any two adjacent layers of the hierarchy, generates a Pareto wage rate distribution with $\beta = \log(s)/\log(d)$; the "span of control" is defined as the number of people in the i-th layer (from the bottom) divided by the number of people in the (i+1)-th layer. See, for example, Simon (1957) or Lydall (1968, pp. 127-133).

8) It is a standard assumption in the theory of optimum income taxation that earned income is an increasing function of the wage rate; the classic article is Mirrlees (1971).

9) If $z=0$, then $D'_z=\phi$, i.e., there is no density function $g: R_+\longrightarrow R_+$ such that $wg(w)$ is decreasing on R_{++}: Let $g: R_+\longrightarrow R_+$ be a function such that $z(w)=wg(w)$ is decreasing on R_{++}. Let $b>0$ be such that $K=z(b)>0$. Then $g(w)\ge K/w$ for all $w\in(0,b]$ and therefore

$$\lim_{a\to 0+}\int_a^b g(w)dw = \infty.$$

Hence, g cannot be a density.

Chapter 2

An Empirical Investigation of the Labour Market
Part I: Estimating Distributions

1. Introduction

In this chapter we explore two distributions, namely the distribution of weekly hours of work and that of gross wage rates in the United Kingdom. More precisely, we estimate density functions. Our data cover the years from 1970 to 1985 and are taken from the Family Expenditure Survey (FES) which is a time-series of cross-sectional data. The size of the annual sample studied is around 7,800 workers.

Density functions are usually estimated under the hypothesis that the functional form of the distribution is known. Since we do not have this knowledge here, we use nonparametric estimation methods. That is, we allow the data to "speak for themselves" in determining the shape of the unknown distribution. We will give examples showing that our results are insensitive to the technical detail of the estimation procedure.

The density of the aggregate labour supply (resp. gross wage rate) distribution was estimated for each odd numbered year of the period under consideration. Using the data of the 1983 FES, we explore the distribution of the two variables within eight subgroups of workers. The picture that emerges has the same feature for other years. We also have a look at the distribution of earned income in the year 1983. The gross wage rate distributions shown in this chapter will be needed in the second part of Chapter 3 in order to compute the elasticity of per capita labour supply (resp. per capita net earnings) with respect to the wage level.

It turns out that the labour supply distributions do not change very much over the years. In fact, when estimating the density functions we did not expect to obtain such a "stable picture". The mean of the data decreases somewhat while its variance increases. However, in the years from

1971 to 1977 we are essentially faced with one distribution. The gross wage rate distributions are considerably less stable. It is interesting to observe that the proportion of individuals receiving a low (resp. high) wage rate increases in the years 1977-85. A brief data analysis shows a steady increase in the sample proportion of female workers and a general switch from manual to non-manual occupations.

Thus far there exists no satisfactory theory of the distribution of wage rates and personal incomes in a market economy. Though there is an enormous literature on this topic, attempts to derive a particular distribution from assumptions about the operating of the economy have always been rather ad hoc. For example, the theory of the lognormal distribution is essentially based on the so-called "law of proportionate effect" due to Gibrat (1931) and an application of the central limit theorem. As M. Friedman (1953) points out, "this absence of a satisfactory theory of the personal distribution of income...is a major gap in modern economic theory" (p. 277).[1)]

Nevertheless, it is a useful exercise to test whether a given sample could have been generated by a density function belonging to a known parametric class of distributions. Since the lognormal distribution has received very much attention in the literature, we tested for all years from 1970 until 1985 and for several groups of workers the hypothesis that the data stem from a lognormal distribution. The variables chosen were: gross wage rate, gross earnings, net earnings and weekly hours of work; the test statistic employed was the Kolmogorov D-statistic. Although the null hypothesis was rejected in most cases, the test results were not as disastrous for the lognormal distribution as one might have suspected. It is interesting to observe that the hypothesis of a lognormal distribution of gross wage rates finds support for full-time workers in non-manual occupations.

The chapter is set up as follows. Section 2 provides a description of the data set. Section 3 introduces nonparametric density estimation methods and applies them to the FES sample of "all workers" for the year 1983; Section 4 investigates eight subsamples. The topic of Section 5 is the stability of the aggregate distributions in the years 1971-85. Finally, Section 6 presents the results of the goodness-of-fit tests. The diagrams relating to Sections 3-5 are always plotted at the end of the corresponding section.

2. The FES Data

The Family Expenditure Survey is a sample survey of the household population in the United Kingdom which provides very detailed information on expenditure patterns as well as on various sources of income and weekly hours of work. The households in the sample vary from year to year, i.e., the FES is a time-series of cross-sectional data and not a panel. The Survey has been in continuous operation since 1957 and is considered as one of the best existing data sets.

Since 1967 the annual set sample (which contains ineligible addresses) has been about 11,000 addresses, of which around 10,750 are selected in Great Britain and 250 in Northern Ireland. The effective sample each year is around 10,000 households and typically some 7,000 households (i.e., approximately 20,000 individuals) agree to participate in the inquiry.

The pre-selected addresses are visited by interviewers, and each member of the household aged 16 (15 before 1973) and over is asked to provide information on both expenditure and income. Only those households where each such person (called a "spender") cooperates are included in the data set. The spenders of a cooperating household receive a small payment. At the preliminary interview a questionnaire covering various sources of income is put individually to each spender. Expenditure information, in turn, is collected partly by interview and partly by diaries which have to be kept over a period of 14 days by the spenders.

The FES in Great Britain is conducted by the Social Survey Division of the Office of Population Censuses and Surveys on behalf of the Department of Employment; the overall design of the sample and its content are kept under review by an inter-departmental committee under the chairmanship of the Central Statistical Office. The FES in Northern Ireland is carried out separately by the Policy, Planning and Research Unit of the Department of

Finance, Stormont. However, both surveys use the same questionnaires and coding instructions and therefore the two data sets can be merged.

The main purpose of the survey was originally to obtain expenditure data for the construction of weights for the retail price indices. However, the FES has become a multi-purpose survey. As the Family Expenditure Handbook remarks, "a number of government departments value the FES solely for its income data" (Kemsley et al., 1980, p. 2). Academic users have access to anonymised FES computer tapes held by the Social Science Research Council Survey Archive at the University of Essex.

In this study we need the FES data on weekly hours of work, gross and net earnings. Clearly, an individual may have more than one job. However, the FES contains only information on hours of work for the *"most remunerative job"* and we therefore focus on main employment. The annual samples underlying our study consist of all workers whose *"last wage/salary from main employment was received last week/month"* (FES code A250). The measures of labour supply provided by the FES for its users are: *"actual/usual weekly hours of work including/excluding paid overtime"*; the measures of labour income are: *"actual/usual weekly gross and net earnings"*. Actual gross earnings are the *"wage/salary, including overtime, bonus, commission or tips the last time the individual was paid"* (FES code 303); actual hours relate to the period for which the individual gave the details of his pay. Since actual earnings and actual hours of work may be subject to substantial temporary variations, we decided to use for individual gross (resp. net) earnings the measure *"usual gross (resp. net) weekly earnings"* and for individual labour supply the measure *"usual weekly hours of work including paid overtime"*. (The FES uses also the terms *"normal earnings"* and *"normal hours of work"*.)

Since these concepts are subjective, there is an instruction on the income schedule of the FES questionnaires for the informant: *"If unable to give usual pay because it varies considerably give average pay received (not basic)"*. It may happen that an individual is unable to state how many hours a week he or she usually works. In this case the household member has to give an explanation: *"If* (this question) *cannot be answered because of the irregular nature of the job give reason"*. For instance, in the year

1979 informants had to answer the following questions: *"How many hours a week do you usually work, excluding meal breaks and overtime?"* (FES code A220); *"on average, how many hours paid overtime do you actually work in a week?"* (FES code A244); and *"what do you usually receive each time you are paid after (before) all deductions?"* (FES codes 329 and 315).

The gross wage rate is obtained by dividing normal weekly gross earnings by normal weekly hours of work. (The income data are recorded on the computer tapes in tenths of pence per week.)

Our data cover the years 1970-85. Throughout this and the next chapter we will assume that the data sets for the years 1971-85 are random samples from the total population of individuals who were in the corresponding year in paid employment (the data for 1970 are possibly unrepresentative; see below). We remark, however, that our contribution may be viewed as a purely descriptive data analysis. It should also be emphasised that users are not uncritical of the FES. The sample size is relatively small; moreover, the FES has a fairly high rate of non-response (in the order of around 30 per cent). In the remainder of this section we comment briefly on the quality of our data.

Interviewers generally ask informants for their pay-slips which are provided by 70-80 per cent of all employees. Furthermore, there is evidence that people consider quite a long period before the date of interview when estimating normal earnings (see Kemsley et al., 1980, p. 71, for details). Nevertheless, it is frequently argued that individuals under-state their income in the FES. In particular, one should have the possibility in mind that there is under-reporting of the earnings of women in part-time employment (Stark, 1978).

Unfortunately, not much is known about the characteristics of non-responding households. One can, however, say the following (see Kemsley et al., 1980, Chapter 10; and Kemsley, 1975). Firstly, the response rate appears to decline with the age of the head of the household. Secondly, households with children appear to show higher response rates than those without. Thirdly, the response rate is not uniformly distributed across regions (the Greater London Area produces the lowest response rate). Finally, the distribution of responses over the year is not completely uniform as well (the number of households responding in December is usually somewhat lower). In other words, non-response is not randomly distributed.

Hence, we cannot exclude that there will be a bias in our estimates due to under-reporting and differential non-response. It appears, however, that one does not have to worry too much about this. For instance, the FES Report 1975 remarks about the survey's earnings data, that these "tend to be slightly deficient, though generally within a few per cent of those indicated by other sources" (p. 3). Two more recent studies reach essentially the same conclusion.

Atkinson and Micklewright (1983) compare various aggregate income data of the 1970-77 FES with the national accounts ("Blue Book") aggregates and summerise their findings by writing: "On the whole the conclusions regarding the reliability of the FES income data are considerably more favourable than those of some earlier investigators...For earnings, the aggregate totals indicate only a small shortfall from the Blue Book total" (p. 50). Atkinson, Micklewright and Stern (1988) provide a detailed comparison of the distribution of earnings and hours of work in the 1971-77 FES with that in the British New Earnings Survey (NES). Like the FES, the NES is a time series of cross-sectional data. There are, however, two important differences. Firstly, the NES obtains its data from employers. Secondly, the sample of the NES is much larger than that of the FES: the size of the annual NES sample is intended to be around one per cent of the employed population.

The authors emphasise in their final assessment that "any divergence may be due to shortcomings of the NES as well as of the FES" and conclude: "The findings with regard to hours and earnings may...be re-assuring to users of both surveys"; in particular, "there is no obvious evidence that the FES figures are seriously affected by higher non-response by those in the upper ranges of the earnings distribution" (Atkinson et al., 1988, pp. 220-221).[2)]

This brief discussion of the Family Expenditure Survey has made much use of the latter two articles and the revised Family Expenditure Survey Handbook (Kemsley et al., 1980). The reader who wants to know more about the survey should consult the FES Handbook. The authors discuss very thoroughly the work of the Social Survey Division and that of the interviewers, sample design aspects and the important issues of non-response and reliability of the FES. Furthermore, the appendix of the Handbook contains the complete questionnaires of the 1979 FES together with other interview documents current in 1979.

The Appendix contains some summary statistics for the labour supply and earnings data in the annual samples of "all workers" and in eight subsamples. All individuals in the samples have positive normal earnings. In a very small number of cases the normal hours of an individual with positive normal earnings are recorded as zero. These individuals were excluded from the samples. Unfortunately, the data of the 1970 FES to which we had access are incomplete; the first 13 weeks of this year are missing.

3. Nonparametric Density Estimation

Let $(x_1,\ldots,x_n)$ be a random sample of real-valued observations from a continuous distribution with probability density ρ. Our aim is to estimate ρ. We assume that we have no information about the probability density beyond $(x_1,\ldots,x_n)$. Hence, we will construct an estimate of ρ (denoted by $\hat{\rho}$) directly from the observations.

In this chapter $(x_1,\ldots,x_n)$ is a sample of workers, and x_i stands for one of the following variables: gross wage rate, weekly hours of work, gross (resp. net) weekly earnings. The diagrams relating to the section are displayed on pages 86-92.

The oldest and most widely used nonparametric estimators of an unknown distribution are the *empirical cumulative distribution function* (cdf) and the *histogram*. The empirical cdf, F_n, of the sample $(x_1,\ldots,x_n)$ is defined by

$$F_n(x) = \frac{1}{n}\cdot\#\{i:\ x_i \leq x\} \quad \text{for all } x.$$

Let [a,b] be an interval somewhat larger than the range of all observations. The histogram is obtained by fixing a partition $a=a_1<a_2<\ldots<a_m=b$ of [a,b] and setting

$$\hat{\rho}(x) = \frac{1}{n(a_{i+1}-a_i)}\cdot\#\{i:\ x_i\ \varepsilon\ [a_i,a_{i+1}[\}$$

whenever $x\ \varepsilon\ [a_i,a_{i+1}[$ $(i=1,\ldots,m-1)$; outside the interval [a,b], $\hat{\rho}(x)$ is set equal to zero.

We remark that none of the data sets considered in this chapter is a simple random sample. However, supposing the observations are the first n in an independent, identically distributed sequence $(x_i:i=1,2,\ldots)$, then the empirical distribution function is an uniformly consistent estimator of the unknown distribution function F. More precisely, by the theorem of Glivenko-Cantelli (e.g., Laha and Rohatgi, 1979, Theorem 2.5.1, p. 114)

$$\lim_{n\to\infty}\sup_{x\in\mathbb{R}}|F_n(x) - F(x)| = 0 \quad \text{with probability 1.}$$

In the case of density estimation the situation is more complicated. Clearly, the shape of the histogram depends on how we have divided the real line into intervals. If the intervals are too large then the fine structure of the data is obscured and hence $\hat{\rho}$ will not be a reasonable estimate of ρ. On the other hand, the histogram becomes unstable if its cells are chosen too small since a shrinking interval contains fewer and fewer observations. This behaviour of the estimator is typical for nonparametric smoothing methods and suggests that $\hat{\rho}$ will be a consistent estimator of ρ only if a_i-a_{i-1} converges "slowly" to zero as the sample size tends to infinity. More precisely, one has to require that a_i-a_{i-1} does not converge as rapidly as n^{-1} to zero (see, e.g, Tapia and Thompson, 1978, Theorem 3, p. 46).

Figure 1 shows the empirical cdf and a histogram of the overall gross wage rate distribution for 1983. (Recall that the FES earnings data are recorded in tenths of pence per week.) Clearly, the cumulative distribution function does not tell us very much about the data. The histogram, on the other hand, immediately reveals that the distribution of gross wage rates is unimodal and skewed to the right. It also appears that there is a "bump" in the central part of the distribution, immediately to the right of the mode, suggesting that the aggregate distribution may be of the form $\rho = \pi_1\rho_1 + \pi_2\rho_2$ (i.e., ρ may be a mixture of two populations).

The histogram is a very useful graphic technique for illustrating the data. From the point of view of density estimation it is, however, not wholly satisfactory. Firstly, it is a step-function and hence exhibits very rapid local variations. In particular, the block form of the estimate causes an unnecessary difficulty if one is not only interested in the unknown density but also in its derivative. (Estimates of ρ' are required in the next chapter.) Secondly, even slight changes in the interval partition may have an effect on the shape of the curve.

We will therefore consider in this section three alternative methods, namely the *kernel method* in Subsection 3.1 and the *spline smoothing* and the *penalised likelihood* approach in Subsection 3.2; the methods will be applied to the FES sample of "all workers 1983". Our aim is to present the ideas behind the methods without going into mathematical details. The mathematically interested reader finds in Prakasa Rao (1983) a comprehensive treatment of the theoretical aspects of nonparametric curve estimation. An excellent non-technical discussion of the subject is given by Silverman (1986). Tapia and Thompson (1978) pay particular attention to the penalised likelihood approach; a standard reference for spline smoothing is Reinsch (1967). Finally, we would recommend to anyone to look through Chapter 24 of the Handbook of Statistics, Vol. 4 (Krisnaiah and Sen, 1984).

3.1. Kernel Estimators

Kernel density estimators belong to the class of the so-called *general weight function estimators*, which are obtained by assigning to each observation x_i a density function $K(\cdot;x_i)$ and setting

$$\text{(W)} \qquad \hat{\rho}(x) = \frac{1}{n}\cdot\Sigma K(x;x_i) \qquad \text{for all } x.$$

(One easily verifies that the histogram is a general weight function estimator.) In order to understand the idea underlying the kernel method, it is useful to begin with a specific member of this class. The question arises how one can free the histogram from a particular choice of subintervals. Consider the density function $\hat{\rho}$ defined by

$$\hat{\rho}(x) = \frac{1}{2nh}\cdot\#\{i: x_i \in]x-h,x+h[\},$$

where h is any positive number; Silverman (1986, p. 12) calls $\hat{\rho}$ the "naive estimator". As we see, the naive estimator is essentially a histogram, but the interval partition has been replaced by intervals]x-h,x+h[($x \in R$). Setting K(x)=1/2 if $|x|<1$, and K(x)=0 otherwise, then $\hat{\rho}$ can be written as

$$\text{(K)} \qquad \hat{\rho}(x) = \frac{1}{n}\cdot\Sigma\frac{1}{h}\cdot K\left(\frac{x-x_i}{h}\right) \qquad \text{for all } x.$$

By generalising (K), one can now define a whole class of density estimators. Notice that the probability density K defined above is symmetric around zero. Let us call a density with this property a *kernel function*. The *(ordinary) kernel density estimator* with *kernel* K and *parameter* h>0 is then defined by (K). The positive number h is called in the literature *smoothing parameter, window width or bandwidth*. One immediately sees that (K) is a special case of (W). Denoting the standard deviation of K by σ and setting $K(x;x_i) = \frac{1}{h}\cdot K(\frac{x-x_i}{h})$, then $K(\cdot;x_i)$ is a symmetric density function with mean x_i and standard deviation $h\sigma$. For instance, if K is the density of the standard normal distribution, then $K(\cdot;x_i)$ is the density of the normal distribution with parameters mean=x_i and variance=h^2. Since (K) is an arithmetic mean of density functions, the kernel estimator is itself a probability density. Obviously, the estimator inherits all the smoothness properties of the kernel function.

For given x, one should think of $K(x;x_i)$ as a *weight* assigned to the observation x_i; the larger h, the more equally distributed will be the weights. For instance, in the case of a "bell-shaped" kernel, $K(x;x_i)$ declines with the distance between x and x_i, but less rapidly for large values of h. In other words, for large values of the smoothing parameter even observations far away from x contribute to the value of $\hat{\rho}$ at x, while for small window widths only observations near to x do so. Consequently, the larger the value of h, the smoother will be the estimator $\hat{\rho}$. The first diagram of Figure 2 shows a kernel estimate of the distribution of gross wage rates; the kernel chosen was the standard normal density (we will discuss the diagrams in the last part of the subsection).

The properties of the estimator do not depend on the technical detail of the kernel function. The symmetry assumption can be weakened, and non-negativity of K is also not really needed. The choice of the window width, however, is crucial for the behaviour of the estimator. Three questions do arise here. Firstly, what is the statistical explanation of the observation that the smoothing parameter determines the shape of the density estimate? Secondly, what is the optimal window width with respect to a given measure for the closeness of the estimator $\hat{\rho}$ to the true density ρ? Thirdly, which values of the window width (even if h is not "optimally" chosen) guarantee that the kernel smoother will be a consistent estimator of the unknown density? There is a large literature on the properties of the kernel method, but for the purpose of this chapter some brief remarks will suffice.

Let us first consider a single point x. A natural local measure of discrepancy is the mean squared error $E\{(\hat{\rho}(x)-\rho(x))^2\}$, where E denotes the expectation operator. The mean squared error can be decomposed into the squared bias and the variance of $\hat{\rho}(x)$. More precisely, mean squared error = $[E\{\hat{\rho}(x)\}-\rho(x)]^2$ + variance of $\hat{\rho}(x)$. One can now say the following: (1) The variance converges to zero as the sample size n tends to infinity; the bias, however, does not depend upon the sample size directly, but only upon the kernel function and the window width. (2) The bias becomes smaller if one decreases the window width while the variance becomes larger. In other words, the systematic error in the estimation of $\rho(x)$ can only be reduced at the expense of increasing the random error, and vice versa, by varying the value of the smoothing parameter.

To obtain a measure for the global accuracy of $\hat{\rho}$ as an estimator of ρ, it is standard practice in the literature to integrate the mean squared error, i.e., one examines the expression

$$\text{IMSE} = \int E\{[\hat{\rho}(x)-\rho(x)]^2\}dx.$$

One can show via a Taylor series expansion of the unknown density ρ that an approximate formula for the optimal window width, from the point of view of minimising IMSE, is given by

$$h^* = Var(K)^{-2/5} \cdot \{\int K(t)^2 dt\}^{1/5} \cdot \{\int \rho''(x)^2 dx\}^{-1/5} \cdot n^{-1/5},$$

where Var(K) denotes the variance of K (see Silverman, 1986, pp. 38-40; and Parzen, 1962, Lemma 4A).

If we set the window width proportional to $n^{-1/5}$ then the (integrated) mean squared error converges to zero at the rate $n^{-4/5}$, and hence the kernel estimator is a consistent estimator of the unknown density. (Recall that an estimator will be consistent if its mean squared error approaches zero.) We remark that in the case of the histogram the IMSE is of the order $n^{-2/3}$ if the cells of the histogram are "optimally" chosen (see, e.g., Tapia and Thompson, p. 48).

The crucial condition ensuring pointwise consistency of the estimator is the following: the window width must converge to zero as the sample size goes to infinity, but not as rapidly as n^{-1}. Clearly, the window width has to decrease in order to reduce the bias. The condition on the rate of convergence, in turn, is equivalent to the requirement that at each point x with $\rho(x)>0$ the expected number of observations falling in the shrinking interval [x-h,x+h] tends to infinity as the sample size becomes larger and larger; this implies that the variance of the kernel estimator converges to zero (the classic article on consistency is Parzen, 1962).

As we see, the optimal window width depends upon the unknown density ρ. Nevertheless, the above formula may be used as a starting point for finding a suitable value of the smoothing parameter. In this study the kernel K is always the density of the standard normal distribution. If the true density ρ is also normal, with standard deviation σ, we obtain

$$\text{(P)} \qquad h^* = 1.06 \cdot \sigma \cdot n^{-1/5}.$$

A quick way of finding a pilot value for h is therefore to estimate σ by the sample standard deviation and to substitute this value into (P). This was done here. In the case of regression estimation [Chapter 3] (P) worked very well. In our opinion, however, the so obtained densities were in general "too smooth" and we therefore decreased the window width somewhat.[3]

The drawback of using a fixed window width across the whole sample is that the estimated density may be unstable in its lower and upper range. We can prevent rapid variations of the estimator in the tails of the distribution by increasing the value of the smoothing parameter. In this case, however, one typically "oversmooths" the central part of the density. The best way of dealing with the problem is to use a larger window width in regions where we have only relatively few observations.

Let $t_1,\ldots,t_n$ be positive numbers. The *adaptive kernel density estimator* with *(global) smoothing parameter* h, *kernel function* K and *local bandwidth factors* $t_1,\ldots,t_n$ is then defined by

$$\text{(A)} \qquad \hat{p}(x) = \frac{1}{n}\cdot\Sigma\frac{1}{t_i h}\cdot K\left(\frac{x-x_i}{t_i h}\right) \qquad \text{for all } x.$$

Notice that the i-th summand in (A) is a density function with mean x_i and standard deviation $t_i \cdot h \cdot$(standard deviation of K). We want to choose the weights $t_1,\ldots,t_n$ in such a manner that t_i is small (resp. large) if there are "many" (resp. "only a few") observations in a neighbourhood of the corresponding data point x_i. The general strategy is to construct the local bandwidth factors from a pilot estimate of the unknown density. In the present study this is done as follows (for a discussion of the method, see Silverman, 1986, pp. 100-110):

(a) start with an ordinary kernel density estimate, $\tilde{p}$, as defined above,

(b) set $g = (\tilde{p}(x_1)\cdots\tilde{p}(x_n))^{1/n}$, and put
$t_i = (g/\tilde{p}(x_i))^{1/2} \quad (i=1,\ldots,n)$.

Because of the factor g in the definition of the t_i, the geometric mean of the local bandwidth factors is equal to one irrespective of the scale of the data. One can therefore use in (A) the same value for h as in the pilot estimate (K). We remark that the above two-stage estimation procedure is insensitive to the mathematical detail of the pilot estimator. In principle, any convenient estimator (e.g., a histogram) can be used to construct the t_i.

In the next chapter we also need an estimator for the first derivative of the unknown density function. Differentiating (K) and (A) with respect to x yields the estimators

$$\hat{\rho}'(x) = \frac{1}{n}\cdot\Sigma\frac{1}{h^2}\cdot K'(\frac{x-x_i}{h}) \qquad \text{(ordinary kernel estimator)}$$

and

$$\hat{\rho}'(x) = \frac{1}{n}\cdot\Sigma(\frac{1}{t_i h})^2\cdot K'(\frac{x-x_i}{t_i h}) \qquad \text{(adaptive kernel estimator).}$$

We now apply the two methods to the FES data. All subsequent kernel estimates were obtained by using the standard Gaussian kernel

$$K(x) = \frac{1}{\sqrt{2\pi}}\cdot\exp\{-\frac{1}{2}\cdot x^2\}, \qquad x\varepsilon R.$$

Figures displaying both an adaptive and an ordinary kernel estimate were produced as follows. Firstly, the same value of the smoothing parameter was used in the two estimations. Secondly, the ordinary kernel estimate served as the pilot estimate required to obtain the adaptive kernel smoother.

In Figure 2 we see kernel estimates of the density of the gross wage rate distribution for the year 1983. Comparison of Figure 2 with Figure 1 shows that one can treat the kernel estimates as if they were smoothed-out histograms. The density estimates for the gross wage rate data are unimodal and skewed to the right. Furthermore, there is a "bump" in the density (immediately to the right of its mode).

A typical feature of the adaptive kernel smoother is that its graph is somewhat more compressed than that of the ordinary kernel smoother. As we see, the local bandwidth factors make the bump in the distribution much more visible. The effect of varying the window width in Figure 2 is as follows. Decreasing h=250 by around 45 per cent transforms the graph of the adaptive kernel smoother into a slightly bimodal density having the shape of the histogram plotted in Figure 1, while the ordinary kernel smoother assumes the shape of the adaptive kernel estimate with h=250, exhibiting, however, some random fluctuations in the upper tail of the distribution. Since the FES sample of "all workers 1983" contains a high proportion of individuals earning a low wage, the density estimates remain stable in the lower range of the distribution even if one selects fairly small values for the window width. If one increases the window width by around 45 per cent, then the bump in the graph of the ordinary kernel smoother disappears, and the adaptive kernel estimate assumes the shape of the ordinary kernel estimate with h=250.

The standard measure for the skewness of a distribution is its third central moment divided by the third power of its standard deviation. The skewness of the empirical gross wage rate distribution for the year 1983 is 3.85. Table 1b in the Appendix contains for all years from 1970 until 1985 the corresponding value of the sample skewness. We do not know whether the high value of 24.23 in 1970 is attributable to the incompleteness of the FES data for that year (as already mentioned in Section 2, we did not have the data for the first thirteen weeks of 1970; in the following calculations we therefore disregarded the year 1970). The arithmetic mean of the sample skewnesses for the years 1971-85 is 5.07, and their standard deviation is 2.64. Thus, the aggregate gross wage rate distributions are strongly skewed to the right.[4)]

We remark that one obtains essentially the same values if one estimates the skewness by computing the third central moment and the standard deviation of the kernel estimates.[5)] This is not surprising. If one is only interested in certain characteristics of a distribution, such as its skewness, then a precise knowledge of the density function is not required. In this case one should estimate the characteristics by sample statistics and not worry about the density function.

Figure 3 presents an adaptive kernel estimate and a histogram of the distribution of weekly hours of work. A problem with the labour supply data is that they are recorded in whole hours. It should be mentioned that the histogram does not cope very well with strongly discretised data. Even a slight change in the interval partition may have an effect on the shape of the curve. In Figure 3 the histogram cells were constructed in such a manner that the j-th cell contains exactly those observations x_i with $x_i=2j-1$ or $x_i=2j$.

The estimates show that the FES sample has a high proportion of part-time workers. More precisely, 24.4 per cent of the individuals in the sample work less than 31 hours per week. It is interesting to observe that over the interval [5,30] these persons are almost uniformly distributed. The skewness in the labour supply data for the year 1983 is -0.34. Table 1b in the Appendix contains the empirical skewnesses for the other years. The mean and the standard deviation of the figures for 1971-85 are -0.43 and 0.12, respectively.

Finally, in Figure 4 we have drawn an adaptive kernel estimate and a histogram of the distribution of gross (resp. net) weekly earnings. Loosely speaking, we see that the tax function transforms a bimodal gross earnings distribution into a unimodal net earnings distribution. On passing from gross earnings to net earnings the mean and the standard deviation of the data are reduced by around 30% and 35%, respectively; the mean of the gross earnings distribution is £116.74 and its standard deviation is £80.07. The empirical skewnesses are 1.67 (for gross earnings) and 1.87 (for net earnings).

Clearly, the earnings distribution is less skewed to the right than the wage rate distribution because of the high proportion of part-time workers in the labour force. We remark that both distributions have the same shape (and a skewness of around 2.5) if we exclude from the sample those individuals working less than 31 hours per week (see also Section 4). At first glance it is somewhat surprising that the gross earnings distribution is not more skewed than the net earnings distribution. However, the British tax system is fairly linear. Let us give a brief description of the tax function.

The income tax schedule is given by an exemption level y_1, an upper bound y_2 and two tax rates, say, t_1 and t_2; $t_1=0.25$ and $t_2=0.40$. A person with gross earnings y has to pay $t_1(y-y_1)$ if $y\varepsilon[y_1,y_2]$; if $y>y_2$, then the individual is charged $t_2(y-y_1)$. For the vast majority of full-time workers the marginal tax rate is 25 per cent. The national insurance system can be represented by a 3-tuple (s,y_3,y_4), where $s=0.09$ and $y_3<y_4$. Individuals earning less than y_3 do not have to make payments. If $y\varepsilon[y_3,y_4]$, then the

contribution is sy; if $y>y_4$, then the contribution is only sy_4. Hence, the tax function can be written as

$$t(y) = t_1\cdot(y-y_1)\cdot 1_{[y_1,y_2]}(y) + t_2\cdot(y-y_1)\cdot 1_{]y_2,\infty[}(y) + sy\cdot 1_{[y_3,y_4]}(y) + sy_4\cdot 1_{]y_4,\infty]}(y), \qquad y\in R_+,$$

where 1_A denotes the indicator function of the set A, i.e., $1_A(y)=1$ if $y\in A$, and $1_A(y)=0$ otherwise.

In Section 4 we will see that the lower (resp. upper) range of the gross wage rate distribution is essentially composed of manual females (resp. non-manual males). In the upper range of the labour supply distribution we have mainly manual males; of course, we will not find many males in its lower range. We now turn to spline smoothing and maximum penalised likelihood estimation.

3.2. Two Alternative Methods

Let F denote the cumulative distribution function of a univariate distribution with density ρ. Then the density can be obtained by differentiating the cumulative distribution function, i.e., we have

$$\rho(x) = F'(x) \qquad \text{for all } x.$$

By the theorem of Glivenko-Cantelli (e.g., Laha and Rohatgi, 1979, Theorem 2.5.1, p. 114), the empirical distribution function F_n is a very good estimator of F. One can therefore construct a density estimate by differentiating a smooth function which approximates F_n. If one is also interested in an estimate of the derivative of the unknown probability density, one has to differentiate the approximating function twice.

The standard approach to this approximation problem is as follows. One chooses a grid $a_1<a_2<\ldots<a_m$, calculates the corresponding function values $F_n(a_1),\ldots,F_n(a_m)$ and then constructs a smooth curve which passes through the points $(a_i,F_n(a_i))$, $i=1,\ldots,m$. That is, one interpolates the data points. If one uses cubic polynomials in each interval $[a_i,a_{i+1}]$, then the aproximating function is called a *cubic spline*. More precisely, a *cubic*

spline interpolant, P, is a function from $[a_1, a_m]$ into R with the following properties:

(1) in each interval $[a_i, a_{i+1}]$ P is a cubic polynomial (in the following denoted by P_i),

(2) $P(a_i) = F_n(a_i)$ for $i=1,\ldots,m$,

(3) $$\left.\begin{aligned} P_{i-1}(x_i) &= P_i(x_i) \\ P'_{i-1}(x_i) &= P'_i(x_i) \\ P''_{i-1}(x_i) &= P''_i(x_i) \end{aligned}\right\} \text{ for } i=2,\ldots,m-1.$$

If the spline satisfies also the condition $P''(a_1) = P''(a_m) = 0$, then it is called a *natural spline* and represents the shape of a *curved ruler* (i.e., a flexible strip of hard rubber or the like) which is forced to match up with the data points in such a way that the ends are left free; the curved ruler takes up a shape which minimises the potential energy. One can show that for any given set of data points a uniquely determined natural spline exists (see, e.g., Burden et al., 1981, pp. 111-113).

In order to actually calculate densities constructed via a cubic spline smoothing of the empirical cumulative distribution function, we used the NAG (Mark 12) library routines E01BAF and E02BCF. The first diagram of Figure 5a shows a spline density estimate for the gross wage rate data. The empirical distribution function was evaluated at the mesh points $a_k = (10{,}000/26)\cdot k$, $k=0,1,\ldots,26$. Whether or not the derivative of a spline interpolant is a reasonable estimate of the unknown probability density ρ is crucially dependent on a proper choice of the grid. In this respect the spline smoothing approach does not differ from histogram estimation. In Figure 5b the empirical distribution function was evaluated at the points $a_k = (10{,}100/26)\cdot k$, $k=0,1,\ldots,26$. We can see how sensitively the shape of the curve reacts to a small change in the grid.

Recall that the histogram will become unstable if one reduces the width of its cells more and more. The same happens to the spline density: The graph of the derivative of a spline interpolant will exhibit rapid local variations if one chooses a very fine grid. For instance, if one partitions the interval [0,10000] into 200 subintervals of equal length and plots the corresponding wage rate density, then one obtains a curve which oscillates so much that one can hardly speak of a probability density. This behaviour of the estimator can be easily explained. The spline interpolant approximates a step-function. The finer we choose the grid, the better will be the approximation. A perfectly smoothed step-function, however, produces a very poor density estimate.

A drawback of the methods discussed so far is that the estimators are derived in a rather ad hoc way from the definition of a density function. If one wants to avoid such an ad hoc definition, one has to explicitly state the aims of the estimation. In parametric statistics this is accomplished by the maximum likelihood method. Recall that a maximum likelihood estimator is implicitly defined as the solution of a maximisation problem. One cannot apply this method directly to nonparametric curve estimation, but there are approaches related to maximum likelihood. To see this, let g by any density function; furthermore, suppose that $(x_1,...,x_n)$ is a simple random sample drawn from g. Then the *likelihood* (i.e., the joint density) of the sample $(x_1,...,x_n)$ is given by

$$\text{(L)} \qquad L(g) = g(x_1)\cdots g(x_n).$$

One is tempted to define the maximum likelihood density estimator as that density function g which maximises (L). However, in the class of all (smooth) density functions this maximisation problem possesses no solution. One can make the likelihood (L) arbitrarily large by taking densities having spikes at the observations and vanishing almost everywhere else, i.e., densities which converge to a sum of Dirac delta-functions.

In parametric statistics this problem does not occur since one places an a priori restriction on the class of admissible densities over which (L) is to be maximised. This is, of course, exactly what we want to avoid in this study. However, a density which looks almost like a sum of delta-functions is a very poor estimate of the unknown probability density. Such a curve would fit the data very well, but it obviously would exhibit too much rapid variation. Thus, we are once again faced with a conflict between "goodness-of-fit" and "smoothness" which has to be quantified.

This, very naturally, leads to incorporating into the likelihood function a term, say $R(g)\geq 0$, which measures in some sense the roughness of the density g under consideration. Let S be the set of all probability densities g for which R(g) is well-defined and finite. The ***penalised log-likelihood*** of a density g in S is now defined as

$$\text{(PL)} \qquad l(g;\alpha) = \sum \log g(x_i) - \alpha \cdot R(g),$$

where α is a positive number; α is called ***smoothing parameter*** and R(g) is called ***roughness penalty.*** The density $\hat{\rho}$ in S which maximises (PL) over S is called ***maximum penalised likelihood density estimator;*** the larger we choose α, the smoother will be $\hat{\rho}$.

We applied to the FES gross wage rate data a version of the penalised likelihood approach due to Scott, Tapia and Thompson (1980). Suppose the unknown density ρ is concentrated on the interval [a,b] and that $\rho(a)=\rho(b)=0$.[6] In order to obtain a computable approximation to the exact maximum penalised likelihood estimator with roughness penalty

$$(*) \qquad R(g) = \int_a^b g''(x)^2 dx,$$

the authors proceed as follows. A positive integer m and a regularly spaced mesh of points $a=a_0<a_1<\ldots<a_m=b$ are chosen. The roughness penalty (*) is approximated by a sum of second differences, and the set S is replaced by the set, say D_m, of all continuous density functions g which are linear over each interval $[a_i,a_{i+1}]$ $(i=0,1,\ldots,m-1)$ and which satisfy $g(x)=0$ if $x \notin\,]a,b[$. This leads to the maximisation problem

$$\text{maximise} \quad \sum \log g(x_i) \quad - \quad \alpha \cdot \sum \{g(a_{k+1}) - 2g(a_k) + g(a_{k-1})\}^2$$

subject to all densities g in D_m, where $a_{-1}=a$ and $a_{m+1}=b$.
The solution of the above optimisation problem is unique and is called the ***discrete maximum penalised likelihood (DMPL) estimator.*** A computer implementation of the method is incorporated in the IMSL programme library, soubroutine NDMPLE.

The DMPL-estimator is relatively robust against variations in the parameter m, i.e., the shape of the estimated density does not change very much if one changes the mesh spacing $a_i - a_{i-1} = (b-a)/m$. The choice of the smoothing parameter α, however, is again crucial for the behaviour of the estimator. Two DMPL-estimates of the gross wage rate density are plotted in Figure 5. In Figure 5a the density was estimated using $\alpha=5\cdot10^{15}$, $a_0=130$, $a_m=10{,}000$ and $m=64$ (all gross wage rates in the 1983 FES are greater than 130); Figure 5b shows a DMPL-estimate with parameter values $\alpha=10^{15}$, $a_0=130$, $a_m=8{,}000$ and $m=64$. If we set $\alpha\geq10^{17}$, then the bump after the mode of the density in Figure 5a disappears and one obtains curves having a very smooth upper tail. Setting $\alpha\leq10^{14}$ yields density functions with rapid local variations; for $\alpha\leq10^{13}$ we obtained curves which were indistinguishable by eye.

Let us conclude this subsection with some remarks. The section begun with the simple histogram estimator. We then pointed out that one should use refinements of the histogram. Among the various smoothing techniques which have been studied in the statistical literature, the method whose properties are probably best understood is (ordinary) kernel estimation. Nevertheless, it is interesting to ask whether the shape of the empirical density depends upon the mathematical detail of the estimator. We therefore have introduced in this subsection the spline smoothing and the penalised likelihood approach. Comparing the histogram, kernel, spline and DMPL estimates with each other, one sees that the similarity of the curves is indeed striking (we remark that a kernel estimation with a "small" window width leads to a density that differs only very slightly from those displayed in Figure 5b).

We also computed spline and DMPL estimates for the distributions of labour income and of hours of work. After some experimentation with the parameter values, essentially the same pictures emerged as those plotted in Figures 3 and 4. We made, however, two observations. Firstly, it turned out that the upper tail of the DMPL-estimator is more unstable than that of the kernel estimator. Secondly, the grid underlying the spline smoothing approach is more difficult to determine than the smoothing parameter of the kernel (resp. DMPL) estimator. In view of this it appears that kernel estimates shown together with histograms serve the purpose of presenting the data best.

In the next section we will have a closer look at the data.

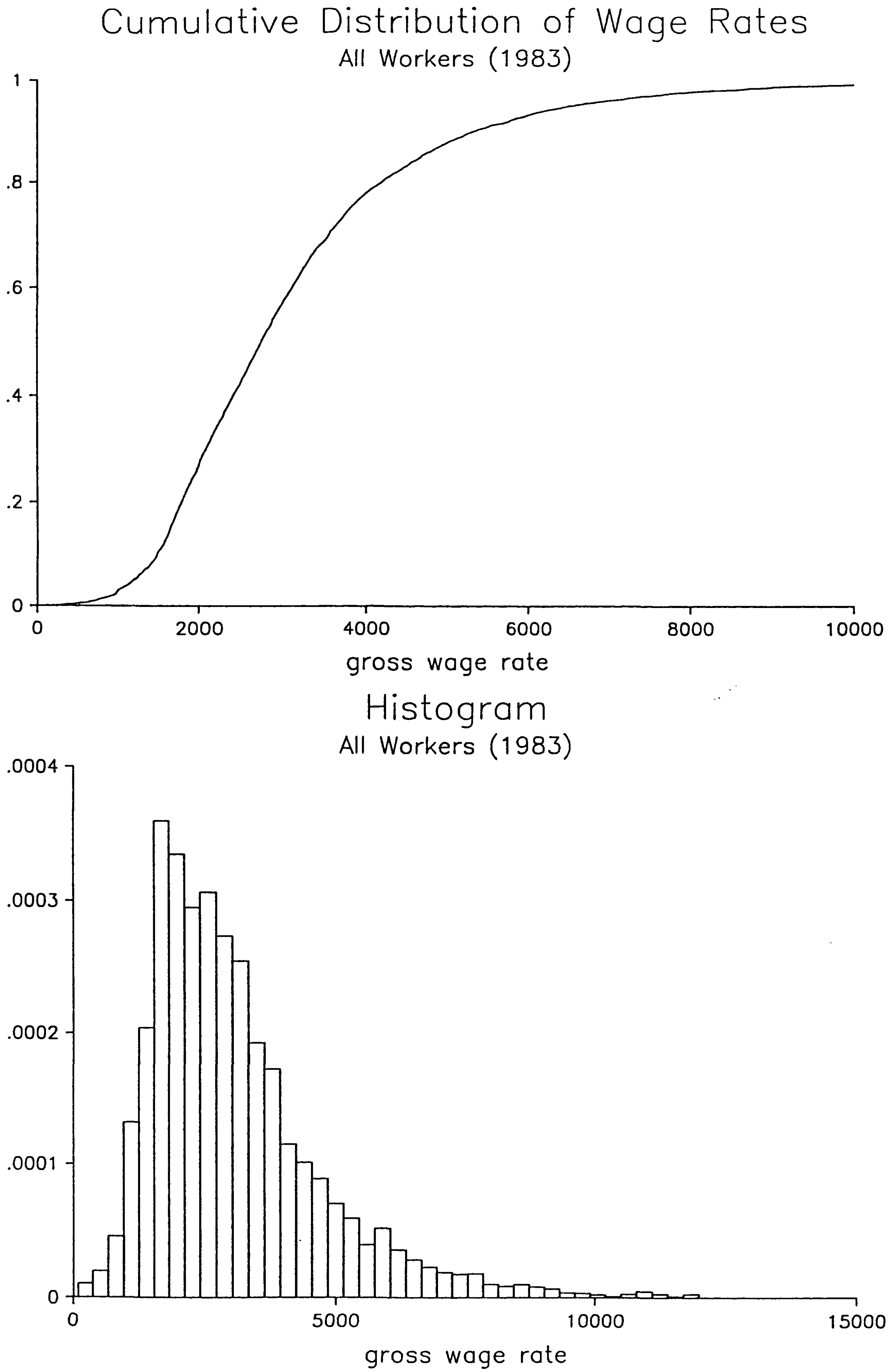
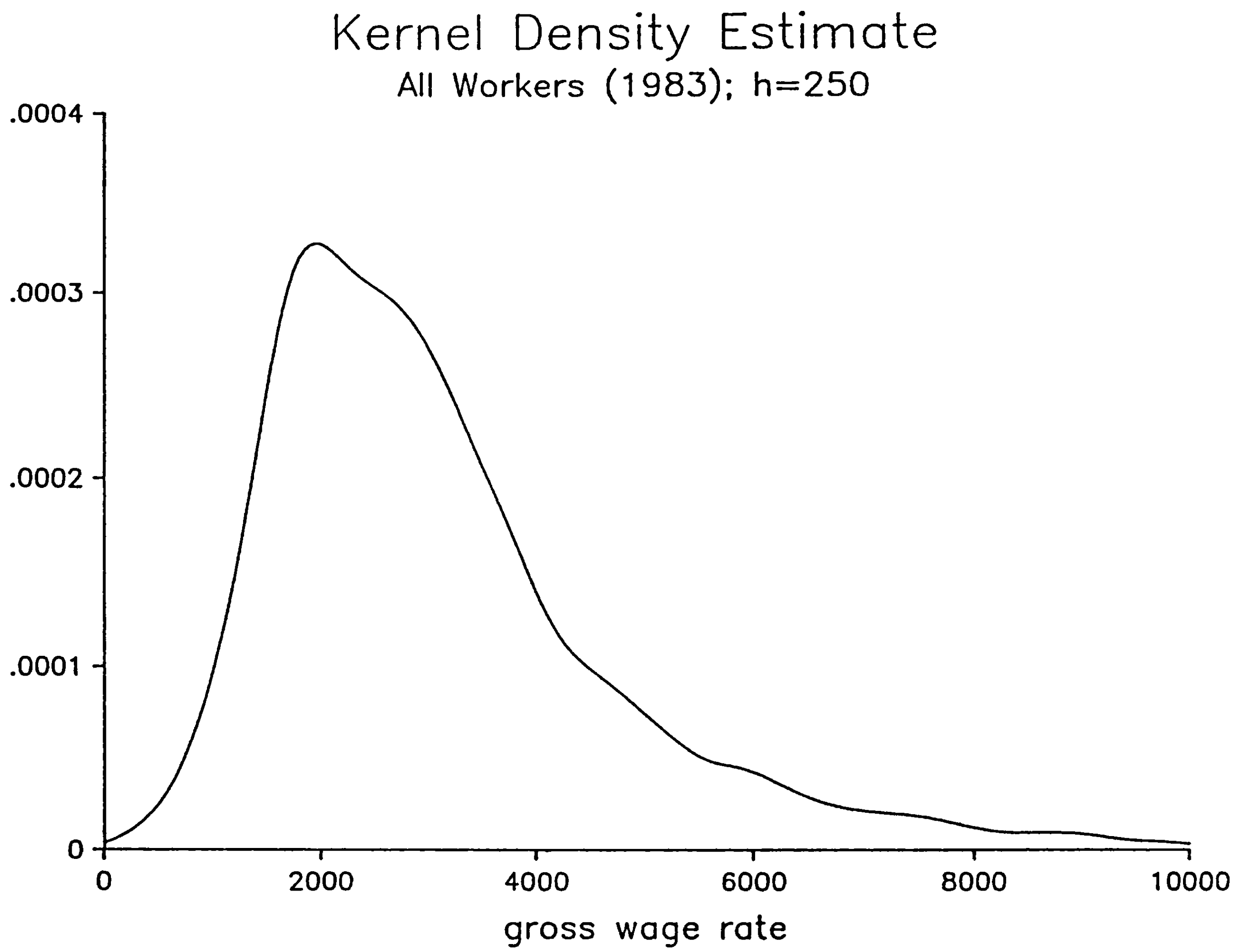

Cumulative Distribution of Wage Rates
All Workers (1983)
1
.8
.6
.4
.2
0
0
2000
4000
6000
8000
10000
gross wage rate
Histogram
All Workers (1983)
.0004
.0003
.0002
.0001
0
0
5000
10000
15000
gross wage rate

Figure 1


Kernel Density Estimate
All Workers (1983); h=250
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate

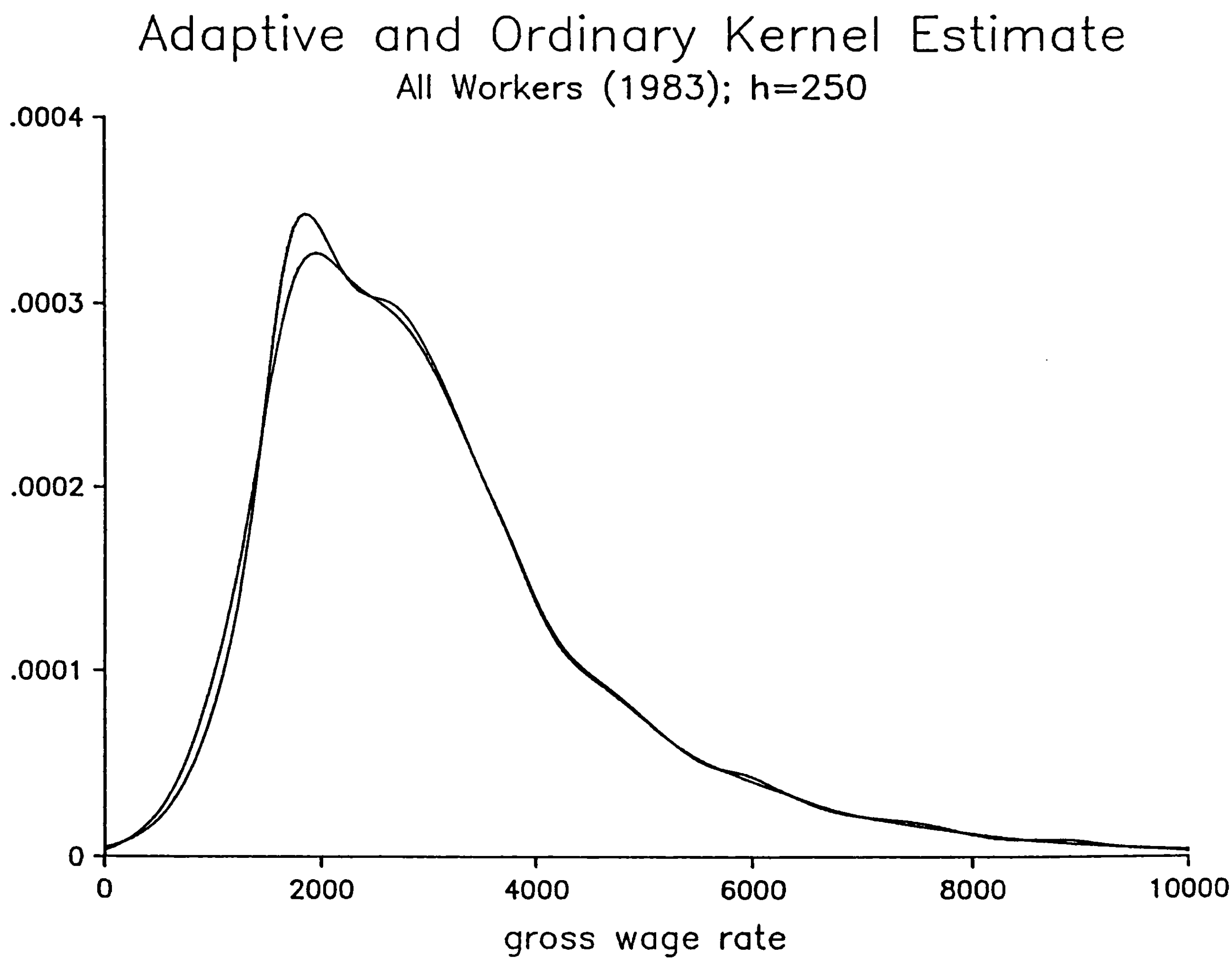
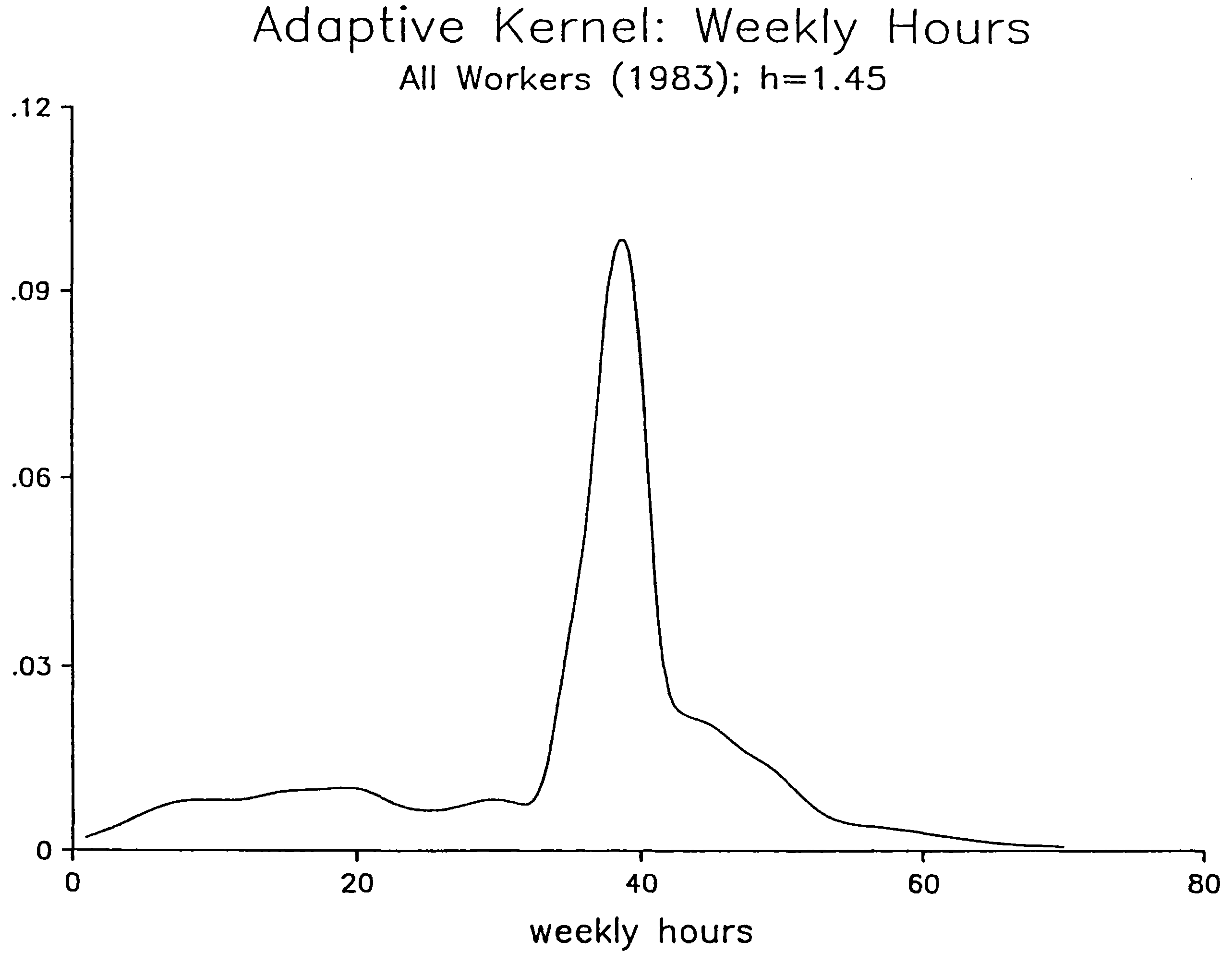

Adaptive and Ordinary Kernel Estimate
All Workers (1983); h=250
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate

Figure 2


Adaptive Kernel: Weekly Hours
All Workers (1983); h=1.45
.12
.09
.06
.03
0
0
20
40
60
80
weekly hours

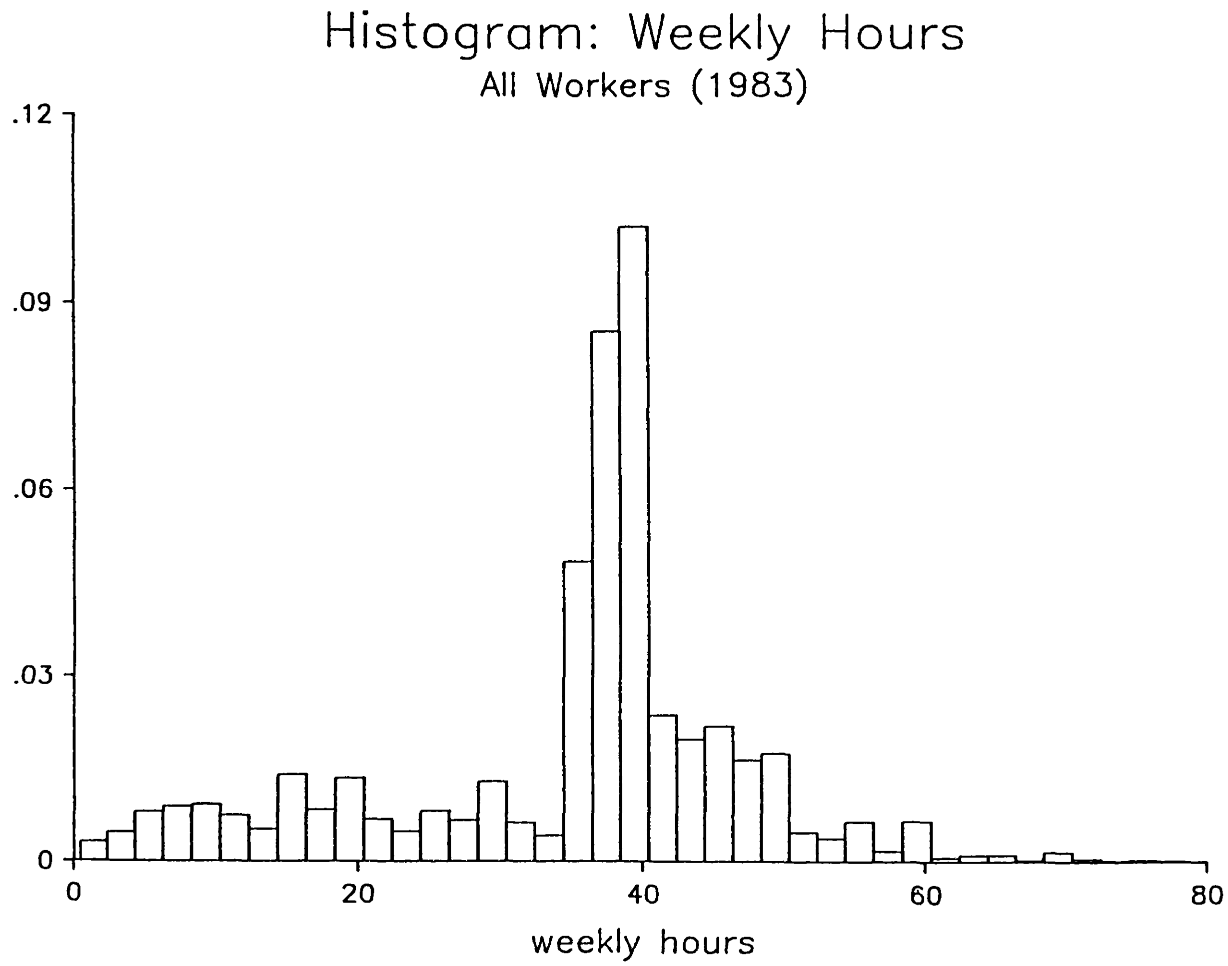

Histogram: Weekly Hours
All Workers (1983)
.12
.09
.06
.03
0
0
20
40
60
80
weekly hours

Figure 3

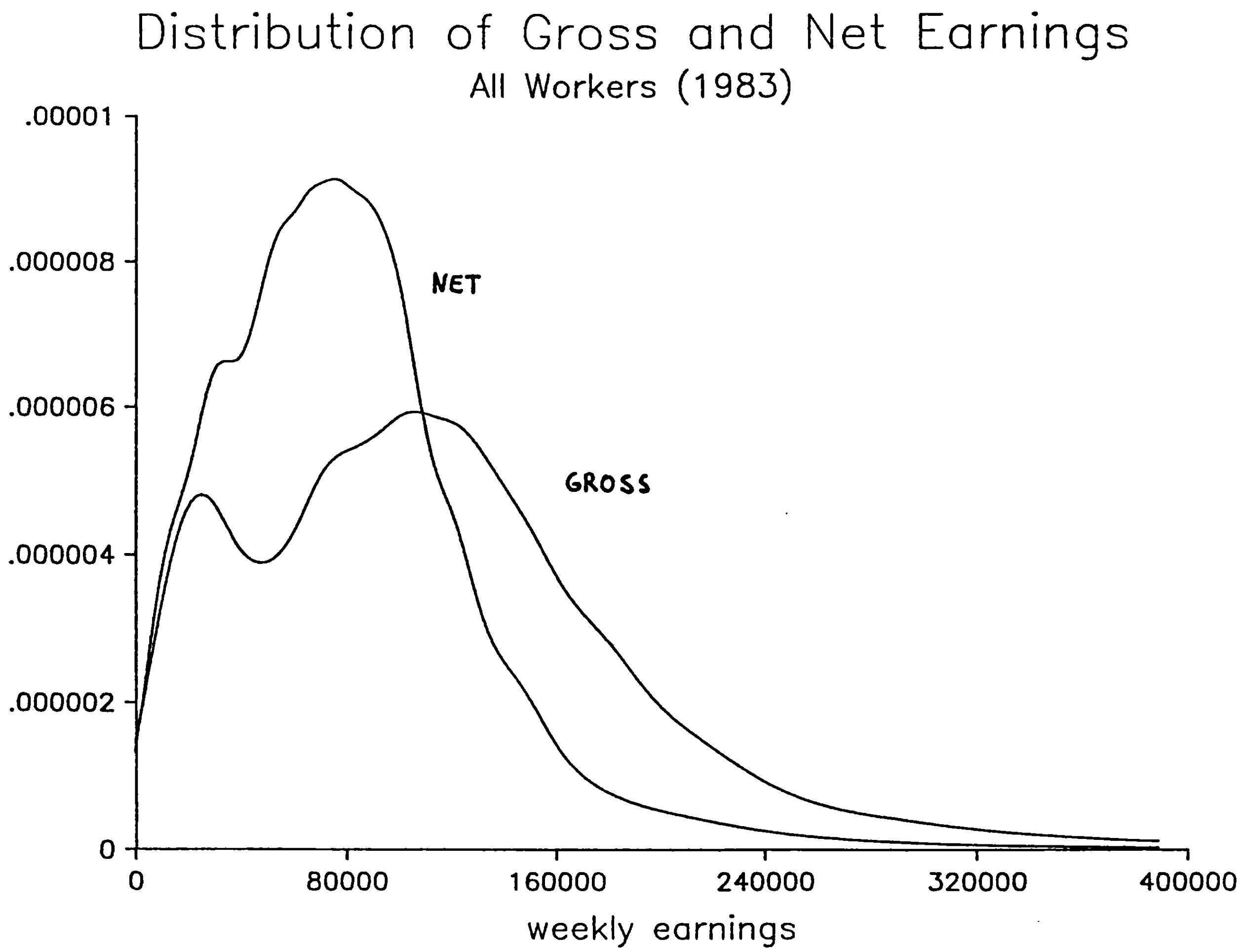
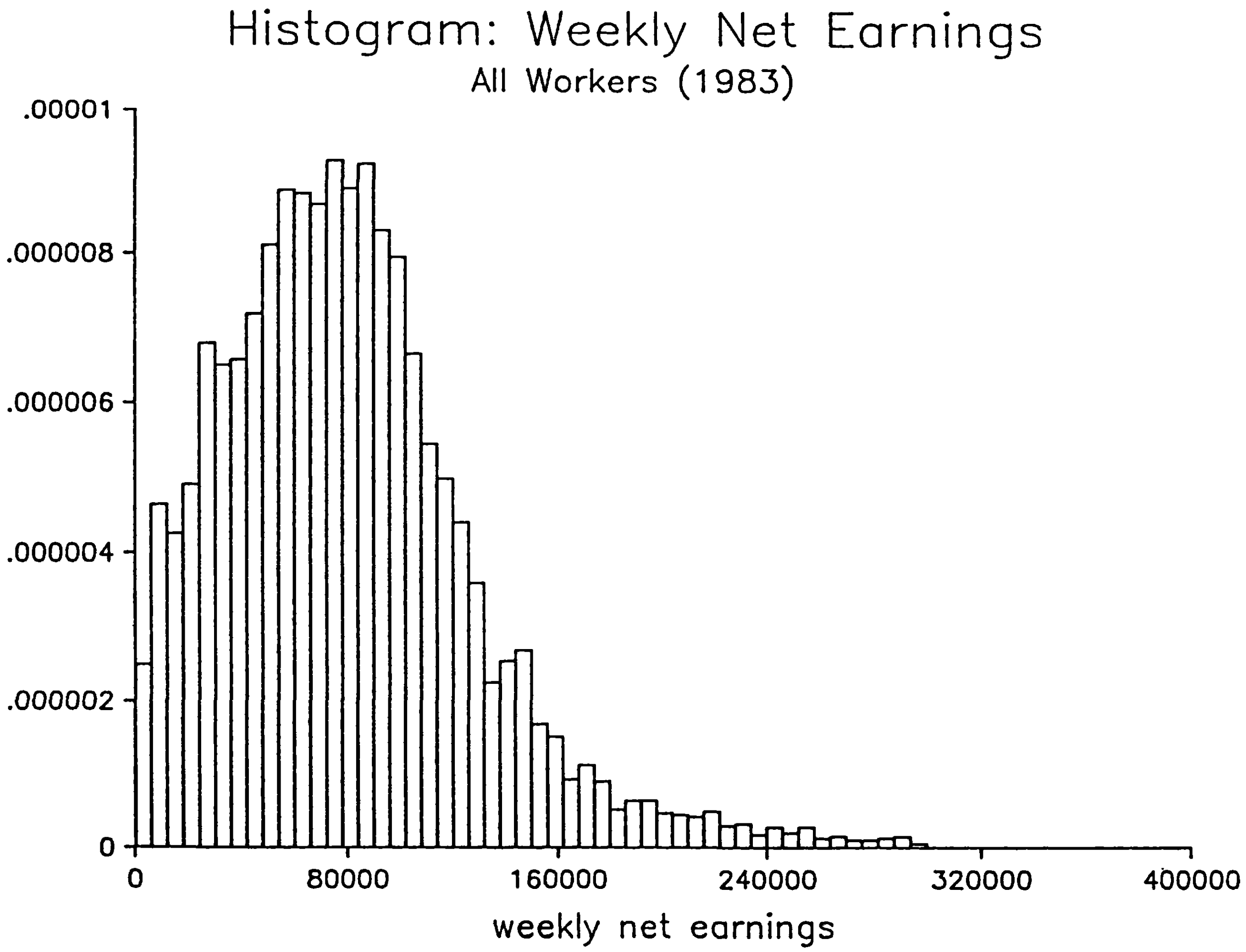


Figure 4a Adaptive kernel estimates


Histogram: Weekly Net Earnings
All Workers (1983)
.00001
.000008
.000006
.000004
.000002
0
0
80000
160000
240000
320000
400000
weekly net earnings

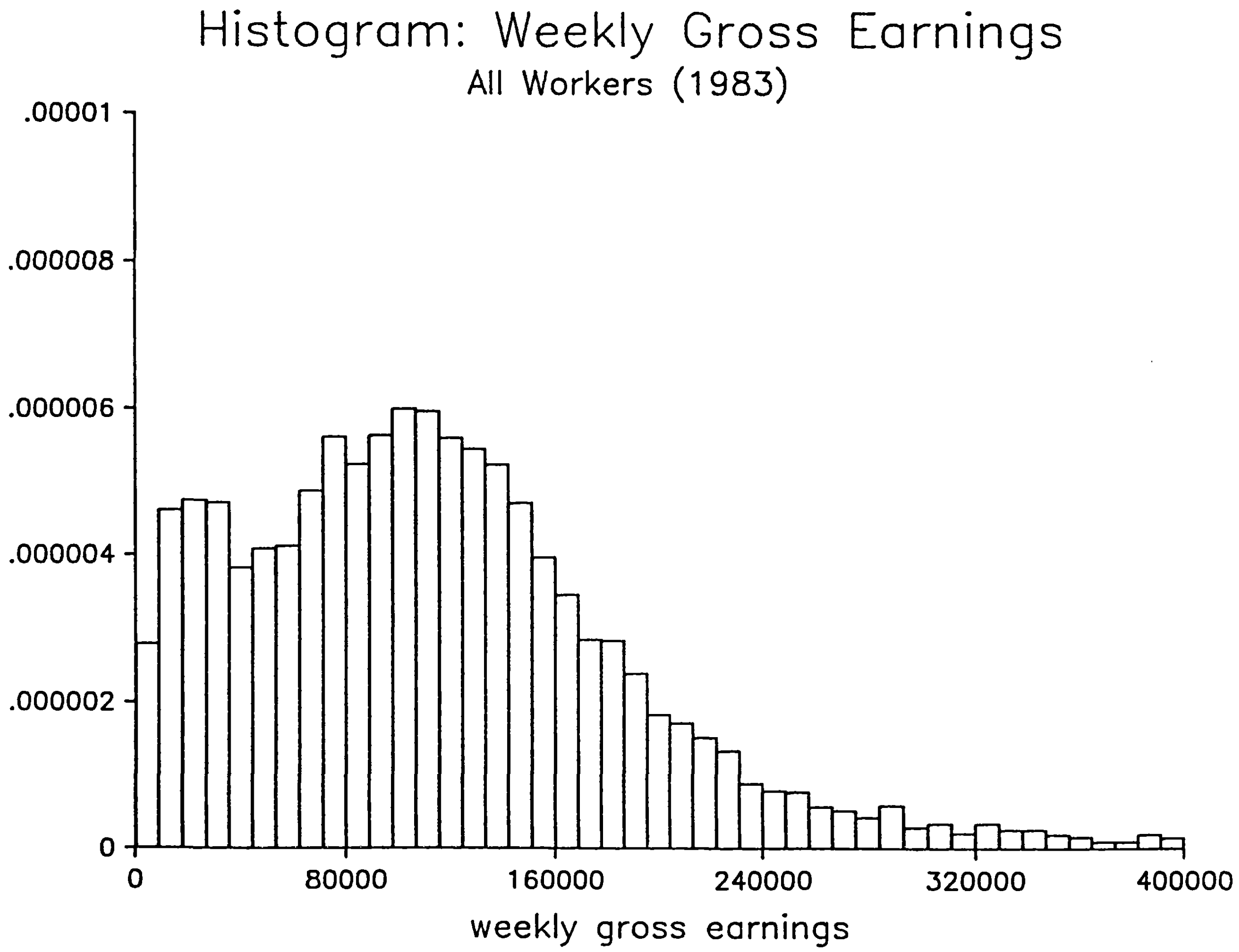

Histogram: Weekly Gross Earnings
All Workers (1983)
.00001
.000008
.000006
.000004
.000002
0
0
80000
160000
240000
320000
400000
weekly gross earnings

Figure 4b

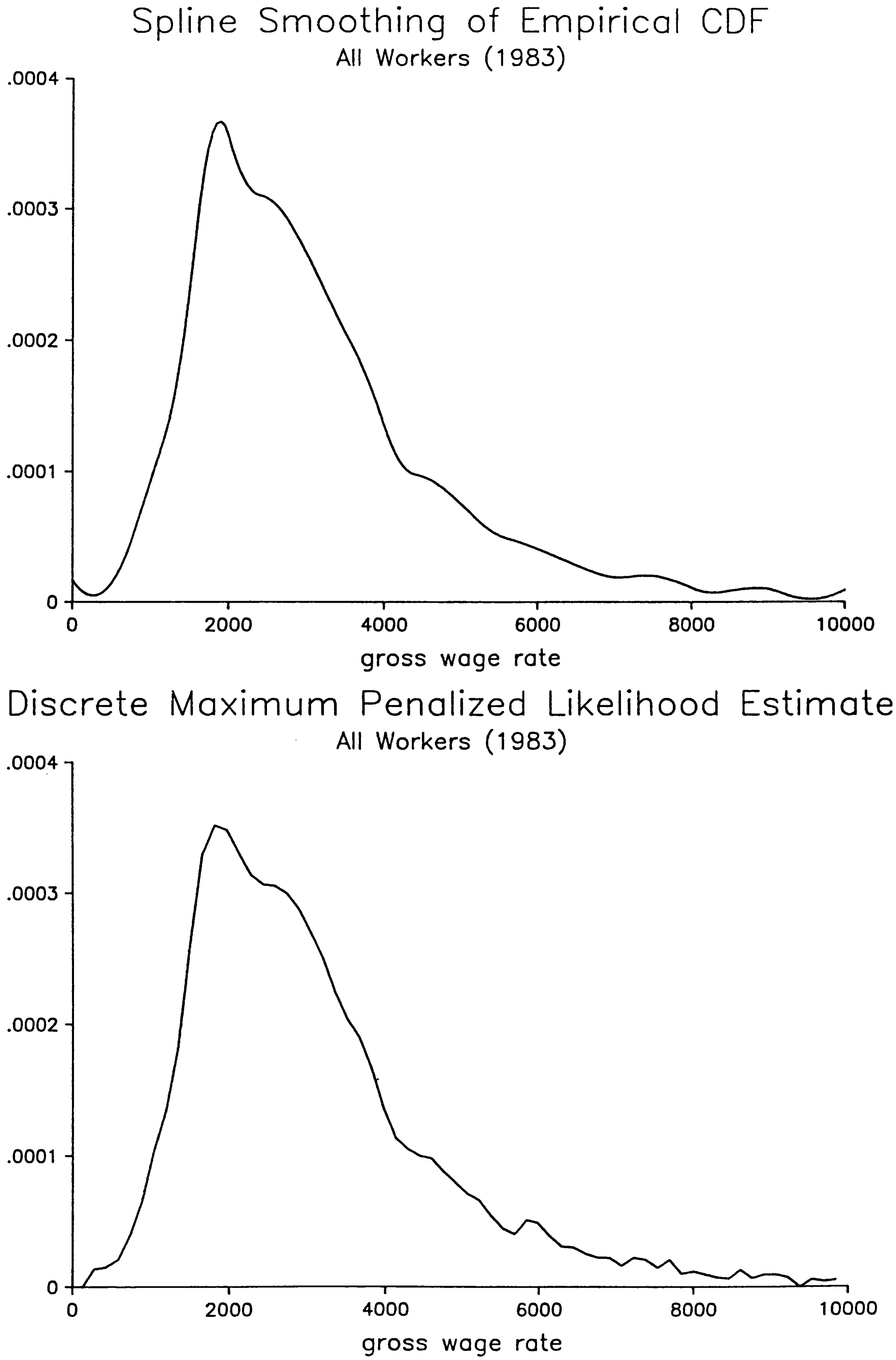
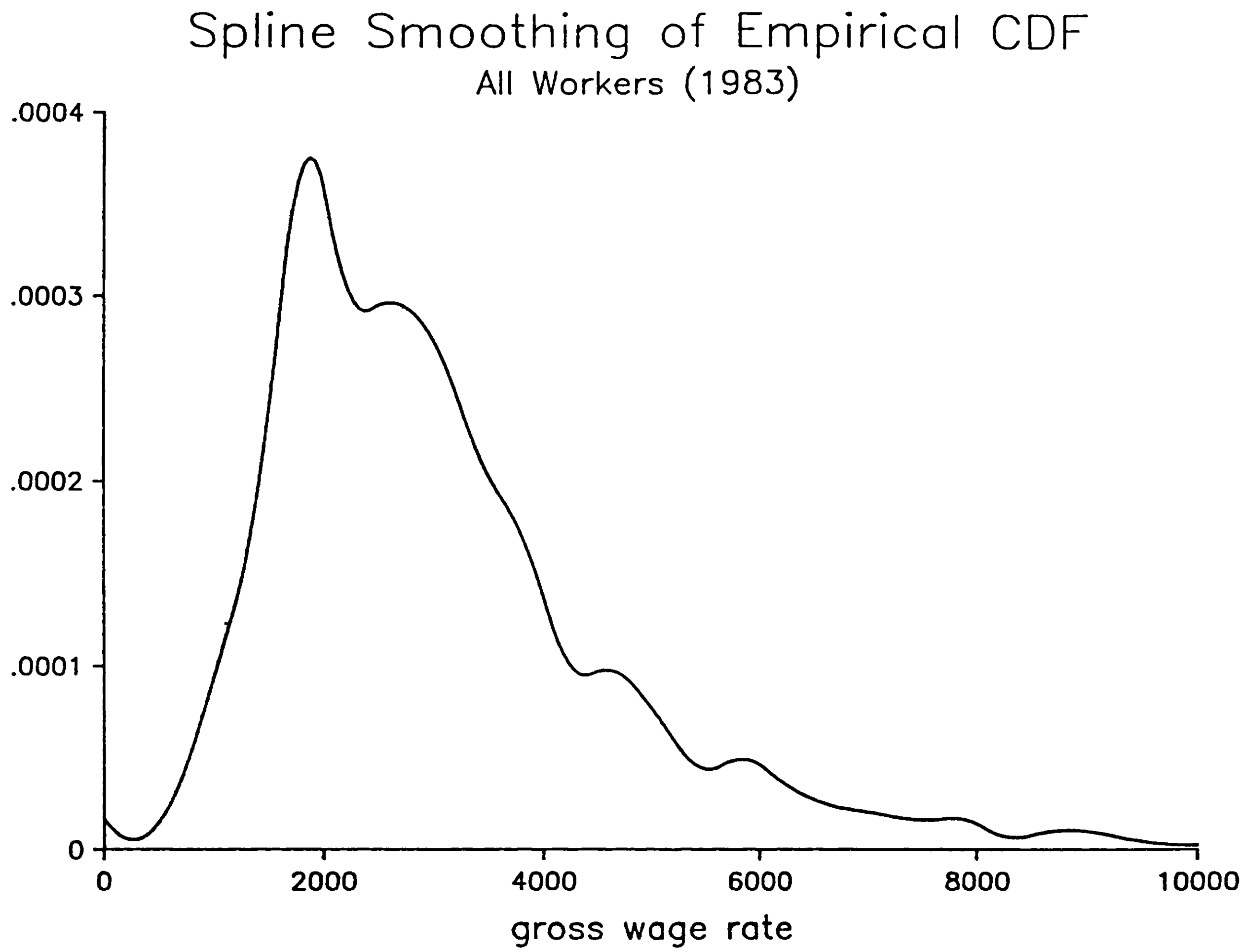

Spline Smoothing of Empirical CDF
All Workers (1983)
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate
Discrete Maximum Penalized Likelihood Estimate
All Workers (1983)
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate

Figure 5a


Spline Smoothing of Empirical CDF
All Workers (1983)
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate

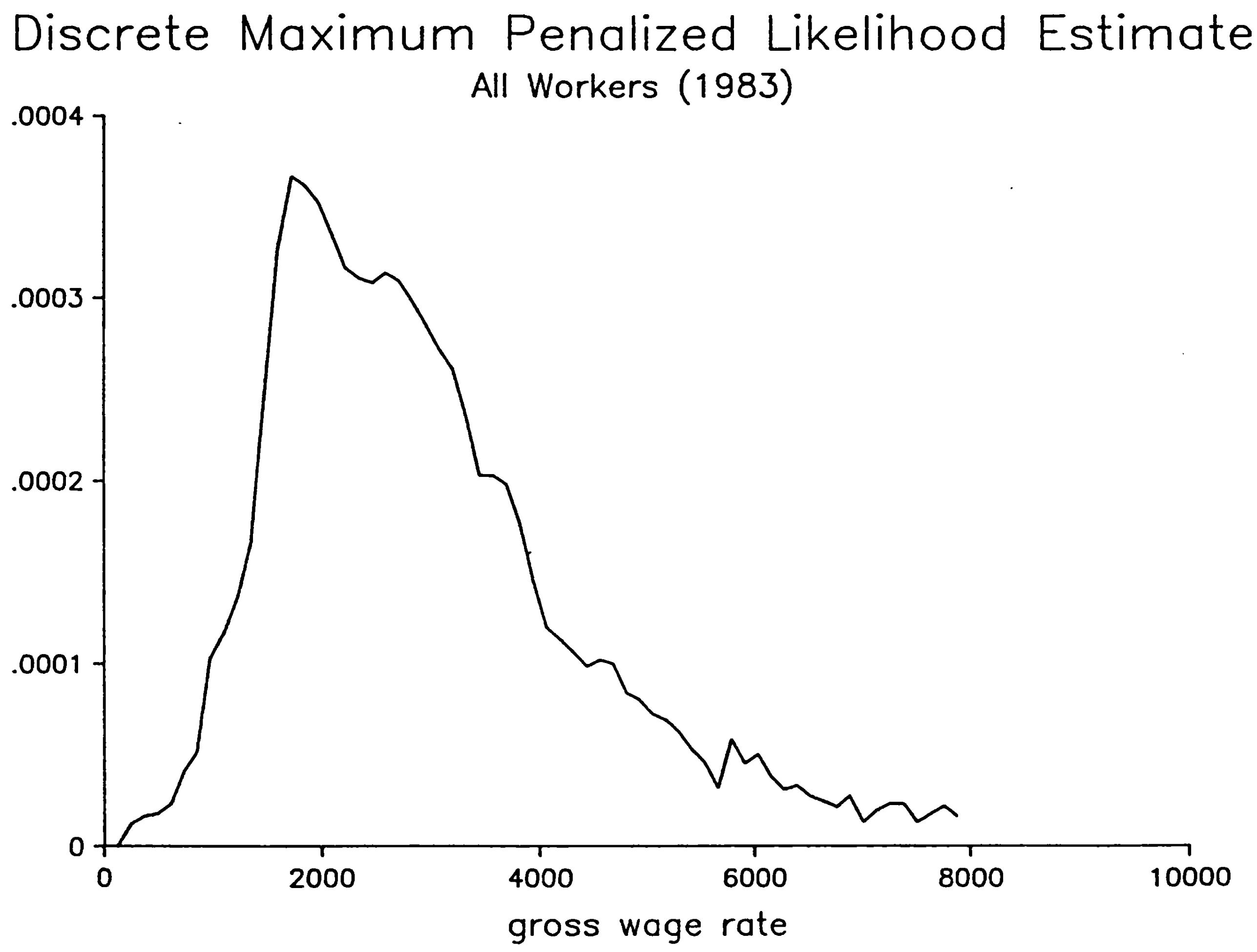

Discrete Maximum Penalized Likelihood Estimate
All Workers (1983)
.0004
.0003
.0002
.0001
0
0
2000
4000
6000
8000
10000
gross wage rate

Figure 5b

4. Investigating Subsamples

The population of "all workers" is a conglomeration of various subpopulations. We will therefore have in this section a look at subsamples of the 1983 FES in order to obtain an insight into the determinants of the densities plotted on the previous pages. There is also another and not less important reason for investigating subsamples. Clearly, postulating that the FES data are realisations of independent and identically distributed random variables would be a heroic assumption (and as we will see in this section, it is not fulfilled either).

On the other hand, this is the standard framework in which one proves theorems on the asymptotic properties of estimators. For instance, in order to show that the (integrated) mean squared error of the kernel estimator is of the order $n^{-4/5}$, one has to postulate that the density estimate is constructed from the first n observations in an independent, identically distributed sequence $(x_i : i=1,2,\ldots)$ drawn from the unknown distribution. If the unknown density ρ is of the form $\rho = n_1\rho_1 + \ldots + n_N\rho_N$, i.e., ρ is a mixture of the densities $\rho_1,\ldots,\rho_N$, then the sample $(x_1,\ldots,x_n)$ falls into N subsamples, say, $(x_i : i\in I_j)$, $j=1,\ldots,N$, where $I_k \cap I_j = \emptyset$ if $k \neq j$, and $I_1+\ldots+I_N = \{1,\ldots,n\}$. Setting $n_j = \# I_j$, we can write [see (K) on p. 74]

$$\hat{\rho}(x) = \Sigma_{j=1,\ldots,N} \frac{n_j}{n} \cdot \left[\frac{1}{n_j} \cdot \Sigma_{i\in I_j} \frac{1}{h} \cdot K\left(\frac{x - x_i}{h}\right)\right].$$

The j-th summand on the right side is a consistent estimator of $n_j \rho_j$ if $h=h(n) \longrightarrow 0$ and $n \cdot h(n) \longrightarrow \infty$ as $n \longrightarrow \infty$. If n is large relative to N, then results on the asymptotic behaviour of the kernel estimator are applicable. In the entire population of workers, however, N may be very large. It should be emphasised that, speaking strictly, an empirical study can only go beyond a descriptive data analysis if the underlying sample can be decomposed into N simple random samples so that N/n is close to zero.

In this section we will investigate the following eight subsamples of the unstratified sample of "all workers 1983": female (resp. male) workers, manual (resp. non-manual) workers, manual female (resp. male) workers, non-manual female (resp. male) workers; the data sets contain both full-time and part-time workers. Of course, the populations from which these samples were drawn are not homogeneous. However, for the purpose of obtaining a better understanding of the aggregate distributions our simple decomposition of the overall population will suffice. In principle, one could have excluded from the data sets individuals working only a few hours per week. We have not done this here since the mere fact that one person works part-time while another is in full-time employment does not necessarily imply that they do not belong to the same "homogeneous" group of workers. A large variance of a statistical variable is in general not an indication of heterogeneity in the underlying population.

Let us now turn to the analysis of the eight samples (resp. populations from which the data stem); an individual will be called *full-time worker* if he or she works usually at least 31 hours per week. We begin with an overview of the diagrams (see pages 102-117). When looking at the wage rate and labour supply distributions, one should keep in mind that the measure for *hours* is subjective (Section 2, pp. 68-69) and that an individual's perception of his *normal hours* may depend on whether the person is a manual or a non-manual worker.

Figure 6 presents the empirical cumulative distribution functions of the gross wage rate distributions for the eight populations. In Figure 7 we see kernel estimates and histograms of the wage rate distributions. The curves drawn in Figures 7a and 7b are the graphs of adaptive kernel estimators; in Figure 7b the window width was decreased by around 35 per cent, i.e., the densities in this figure "follow the data" more closely than those of Figure 7a. In Figure 7c each adaptive kernel smoother is plotted together with the ordinary kernel estimate which was used to construct the local bandwidth factors; the values for the window width are the same as those used in Figure 7b.

The graphs of the pilot estimates have the same shape as those of the adaptive kernel smoothers plotted in Figure 7a. In our opinion the values of the smoothing parameter used in Figure 7b are too small. But the reader may decide himself which diagrams he prefers, for instance, by comparing the kernel estimates with the histograms in Figure 7d. In Figure 8 adaptive kernel estimates and histograms of the distributions of weekly hours of work are displayed. The histograms were constructed in the same manner as the histogram in Figure 3. Finally, Figure 9 shows adaptive kernel estimates and histograms of the distribution of gross (resp. net) weekly earnings within the eight populations.

Notice that the empirical distribution function of the gross wage rate distribution for males (resp. non-manuals) lies everywhere above that for females (resp. manuals); the empirical distribution function for non-manual male (resp. female) workers lies everywhere above that for manual male (resp. female) workers. The gross wage rate densities are unimodal. The skewness in the data is in all cases positive. However, the density for the subgroup "male manual workers" is in its main part essentially symmetric, and the density for the subgroup "female manual workers" is also fairly symmetric (but less than that for male manuals).

We are now able to explain the bump in the aggregate wage rate density (Figure 2). Looking at Figure 7a, we see that the density for the whole population is a mixture of the densities for manuals and non-manuals. The gross wage rate distribution for manual (resp. non-manual) workers is, in turn, a mixture of the distributions for female manual (resp. non-manual) workers and male manual (resp. non-manual) workers. Just as the density for the total population, the density for the subgroup "manual workers" has a bump immediately after its mode. Loosely speaking, the cause of this bump is the female-male pay gap: the density for the subgroup "female manual workers" has a much higher peak than the density for the subgroup "male manual workers", and the former density assumes its maximum at a lower gross wage rate than the latter.

We also estimated wage rate distributions for populations of full-time workers, i.e., we excluded from the sample of "all workers 1983" and from the eight subsamples those individuals who stated in the FES questionnaires that they worked usually less than 31 hours per week. Excluding part-time workers from a sample does not change the shape of the empirical density function very much. All nine densities are unimodal and have no bumps. The distributions are, however, less skewed to the right than those estimated on the samples which contain also part-time workers. The positive skewness of the gross wage rate densities for the subgroups "full-time manual female (resp. male) workers" and "all full-time manual workers" is almost not visible; these three densities are essentially symmetric.

A brief glance at Figure 8 shows that the vast majority of males are full-time workers, while about every second woman works part-time. Thus the lower range of the aggregate distribution displayed in Figure 3 is explained by female labour supply. The labour supply data are with three exceptions skewed to the left; the subgroups with positive sample skewness are: "manual females", "non-manual males" and "all males" (see below).

Since females earn lower wage rates than males and also work less hours, we obtain a bimodal distribution of gross earnings for the total population of workers (Figure 4).

Looking at the earnings data, we see that the densities for males, manual males and non-manual males are unimodal and "bell-shaped". The densities for female workers and the two subgroups are of a very different type. Because of the high proportion of part-time workers among the females, none of these densities is unimodal. All six distributions possess a mode near to the left endpoint of the earnings range. The gross (resp. net) earnings density for the subgroup "female manual workers" shows a three-mode-shape while the other densities are bimodal. Reasonable variations of the window width do not change the multimodality of the density functions.

Since the British tax system bears a strong resemblance to the simple linear tax function, the gross earnings distributions are not more skewed than the net earnings distributions (the only exception is the subgroup "female manual workers"). The tax function shifts the earnings distribution for males to the origin; the result is an almost unimodal aggregate net earnings distribution.

In the remainder of this section we will have a look at the composition of the sample of "all workers 1983" and at four simple characteristics of the distributions. The data set, which provides information on 6833 workers, contains more non-manuals than manuals and more males than females. Contrary to male labour supply, where more individuals work as manuals than as non-manuals, only every third woman is a manual worker. As we have already seen in Figure 8, the vast majority of males work full-time, while around every second woman is in a part-time employment. Part-time male labour is equally divided between manual and non-manual work. While the majority of full-time females can be found in non-manual occupations, part-time female labour is almost equally divided between manual and non-manual occupations; but among the part-time females there are also more non-manuals than manuals.

The values of the sample statistics are taken from the Appendix; CV is an abbreviation for coefficient of variation; the statistics for the earnings data are given in £. Looking at the the sample of "all workers 1983", we have for the four distributions under consideration:

	mean	standard deviation	CV	skewness
gross wage rates	3.18	1.98	0.623	3.85
gross earnings	116.74	80.07	0.686	1.67
net earnings	81.90	52.12	0.636	1.87
weekly hours	35.57	13.08	0.368	-0.34

We now give a brief description of the eight populations.

Female and Male Workers:

43.8 per cent of all workers are females; 48.6 per cent of the females work part-time, while only 5.5 per cent of the males are in part-time employment. Males do not only work more hours per week than females but they also receive considerably higher hourly earnings; the two wage rate distributions differ, however, only very slightly with respect to their skewness. The distribution of weekly hours of work for males is slightly skewed to the right; the skewness in the sample for females is of the same order but has a negative sign. Because of the high proportion of part-time workers the distribution of earned income for females is less skewed to the right than that for males. For females we have the following figures:

	mean	standard deviation	CV	skewness
gross wage rates	2.50	1.49	0.596	3.74
gross earnings	71.07	51.29	0.722	1.39
net earnings	52.27	32.59	0.623	1.53
weekly hours	27.92	12.59	0.451	-0.22

The values of the summary statistics for male workers are:

	mean	standard deviation	CV	skewness
gross wage rates	3.72	2.14	0.575	4.06
gross earnings	152.41	80.39	0.527	1.93
net earnings	105.03	52.80	0.503	2.16
weekly hours	41.55	9.96	0.240	0.20

Observe that the deductions from gross earnings imply for females a decrease in mean earnings of around 26 per cent, while we have for males a reduction of about 31 per cent. This difference follows from the fact that many part-time females have earnings below the exemption level of the income tax (see the brief discussion of the British tax schedule at the end of Subsection 3.1). Notice also that on passing from gross earnings to net earnings we have a slightly larger reduction in the sample standard deviation for females than in that for males. The spread of the data around its mean is reduced in the former group by 36 per cent and in the latter by 34 per cent. It would be interesting to examine how much of the earnings differential between females and males can be attributed to discrimination as opposed to differences in the quality of labour supplied. Unfortunately, the FES lacks information on the level of education and past work experience which one would need for such a study.[7)]

Female Manual and Non-Manual Workers:

35.4 per cent of the females are manual workers. The proportions of full-time workers among female manuals and female non-manuals are 37.7 per cent and 58.9 per cent, respectively; hence 45.5 (resp. 26) per cent of the part-time (resp. full-time) females are manuals. Many of the manual females work for very low hourly wages. Non-manual females earn on average higher gross wage rates than manual females and they also work more hours per week. The gross wage rate distribution is considerably more skewed in the population of "non-manual female workers" than in that of "manual female workers". The distribution of weekly hours of work is slightly skewed to the right for manual females and skewed to the left for non-manual females. Clearly, the positive skewness of the former distribution is a consequence of the fact that among the manuals only every third woman works full-time. The values of the sample statistics for manual females are:

	mean	standard deviation	CV	skewness
gross wage rates	1.85	0.66	0.357	1.58
gross earnings	47.62	34.51	0.725	1.18
net earnings	37.49	22.56	0.602	0.86
weekly hours	24.66	13.25	0.537	0.11

For non-manual females we have:

	mean	standard deviation	CV	skewness
gross wage rates	2.85	1.69	0.593	3.42
gross earnings	83.95	54.35	0.647	1.25
net earnings	60.39	34.36	0.569	1.52
weekly hours	29.70	11.84	0.399	-0.36

Notice that taxes and social security contributions reduce the mean of the earnings data for non-manual females by around 28 per cent, while we have for manual females a decrease in mean earnings of only 21 per cent; as in the case of *female and male workers* the standard deviations are reduced much more equally, namely by 35 per cent for manual females and by 37 per cent for non-manual females.

Male Manual and Non-Manual Workers:

55.7 per cent of all males work as manuals, and 95 per cent of the manual workers are in full-time employment; in the sample of non-manual males 93.8 per cent of the individuals work full-time; hence 50 (resp. 56) per cent of the part-time (resp. full-time) males are manuals. The distribution of weekly hours of work for non-manual males is skewed to the right and that for manual males is skewed to the left. Manual males spend on average more time at work than males in non-manual occupations, but they have substantially lower (hourly) earnings. The distributions of weekly and of hourly earnings for manual (resp. non-manual) males are almost symmetric (resp. skewed to the right). Comparing the earnings data before taxation with those after all deductions, we observe for non-manual (resp. manual) males a reduction in the sample mean of around 32 (resp. 30) per cent and a reduction in the sample standard deviation of around 34 (resp. 35) per cent. The values of the sample statistics for manual male workers are:

	mean	standard deviation	CV	skewness
gross wage rates	2.98	1.15	0.386	1.48
gross earnings	127.70	54.47	0.472	0.81
net earnings	89.17	35.25	0.395	0.88
weekly hours	42.36	9.56	0.226	-0.54

The figures for non-manual male workers are:

	mean	standard deviation	CV	skewness
gross wage rates	4.65	2.67	0.574	3.70
gross earnings	183.54	95.57	0.521	1.74
net earnings	125.01	63.40	0.507	1.98
weekly hours	40.53	10.37	0.256	0.98

Manual and Non-Manual Workers:

46.8 per cent of all workers are manuals; 33.2 per cent of the manuals and 53.2 per cent of the non-manuals are females. Most workers in the two samples work full-time, namely 76 per cent of the manuals and 75.2 per cent of the non-manuals. In the subsample of full-time (resp. part-time) manual workers the proportion of women is 16.4 (resp. 86.2) per cent; in the subsample of full-time (resp. part-time) non-manual workers the proportion of women is 41.7 (resp. 88.2) per cent. As a consequence of the previous figures, manual workers have substantially lower (hourly) earnings than non-manual workers. In both populations labour supply is skewed to the left, but for non-manual workers the value of the empirical skewness is close to zero. The remaining distributions are more skewed to the right for non-manuals than for manuals. The deductions from gross earnings reduce the standard deviations in the two samples by some 35 per cent. The arithmetic means of the data are reduced by around 29 per cent (for manual workers) and 30 per cent (for non-manual workers). The values of the summary statistics for manual workers are as follows:

	mean	standard deviation	CV	skewness
gross wage rates	2.60	1.15	0.442	1.47
gross earnings	101.13	61.63	0.609	0.68
net earnings	72.03	39.89	0.554	0.69
weekly hours	36.49	13.74	0.377	-0.63

For non-manual workers we have:

	mean	standard deviation	CV	skewness
gross wage rates	3.69	2.38	0.645	3.53
gross earnings	130.51	91.16	0.698	1.70
net earnings	90.90	59.56	0.655	1.93
weekly hours	34.77	12.41	0.357	-0.03

The tables in the Appendix show that the picture that emerged for 1983 has the same feature in other years. The reader may find it interesting to look through Routh (1980) who explores the trend in occupational earnings differentials in Great Britain over the period 1906-79.

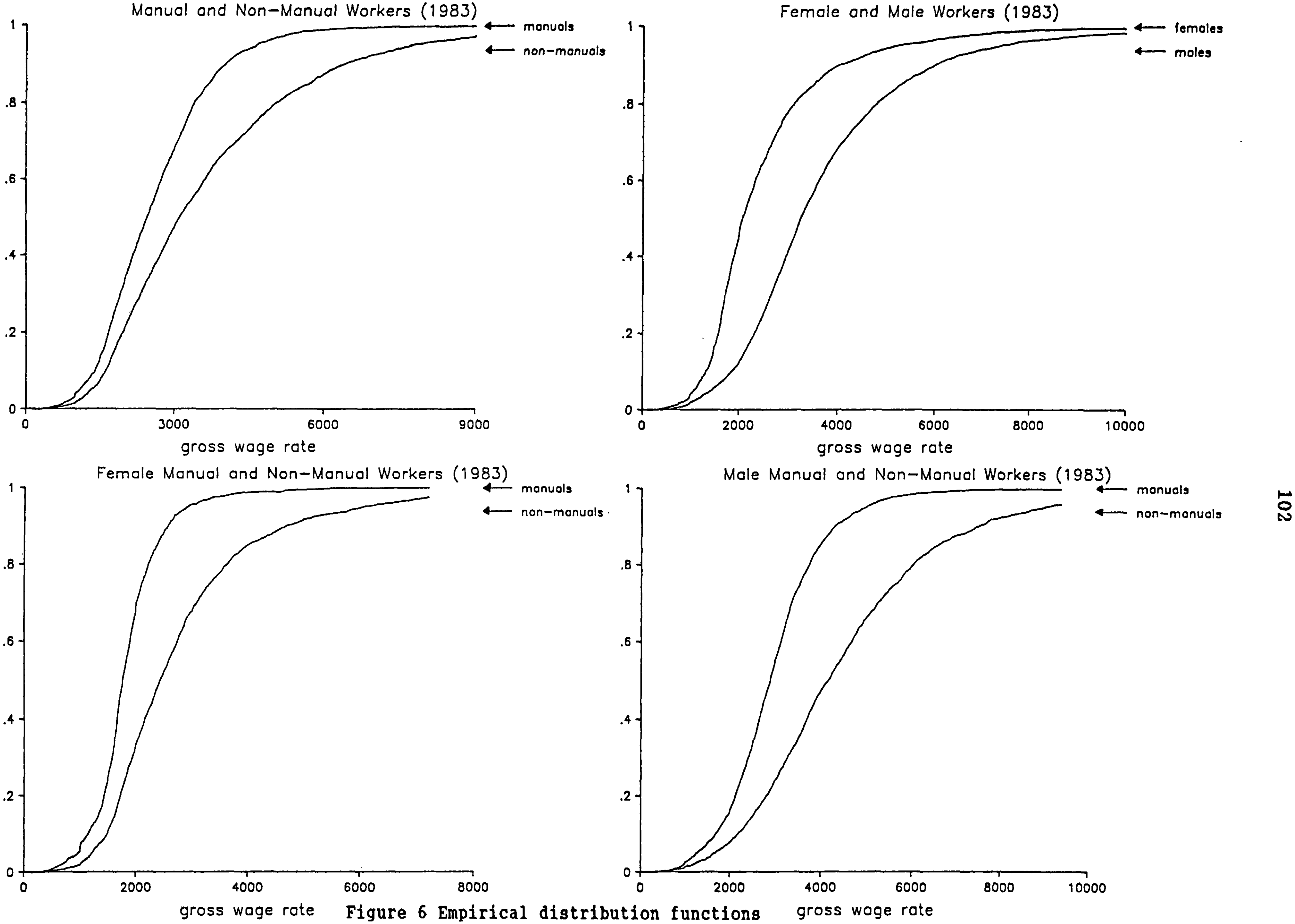


Figure 6 Empirical distribution functions

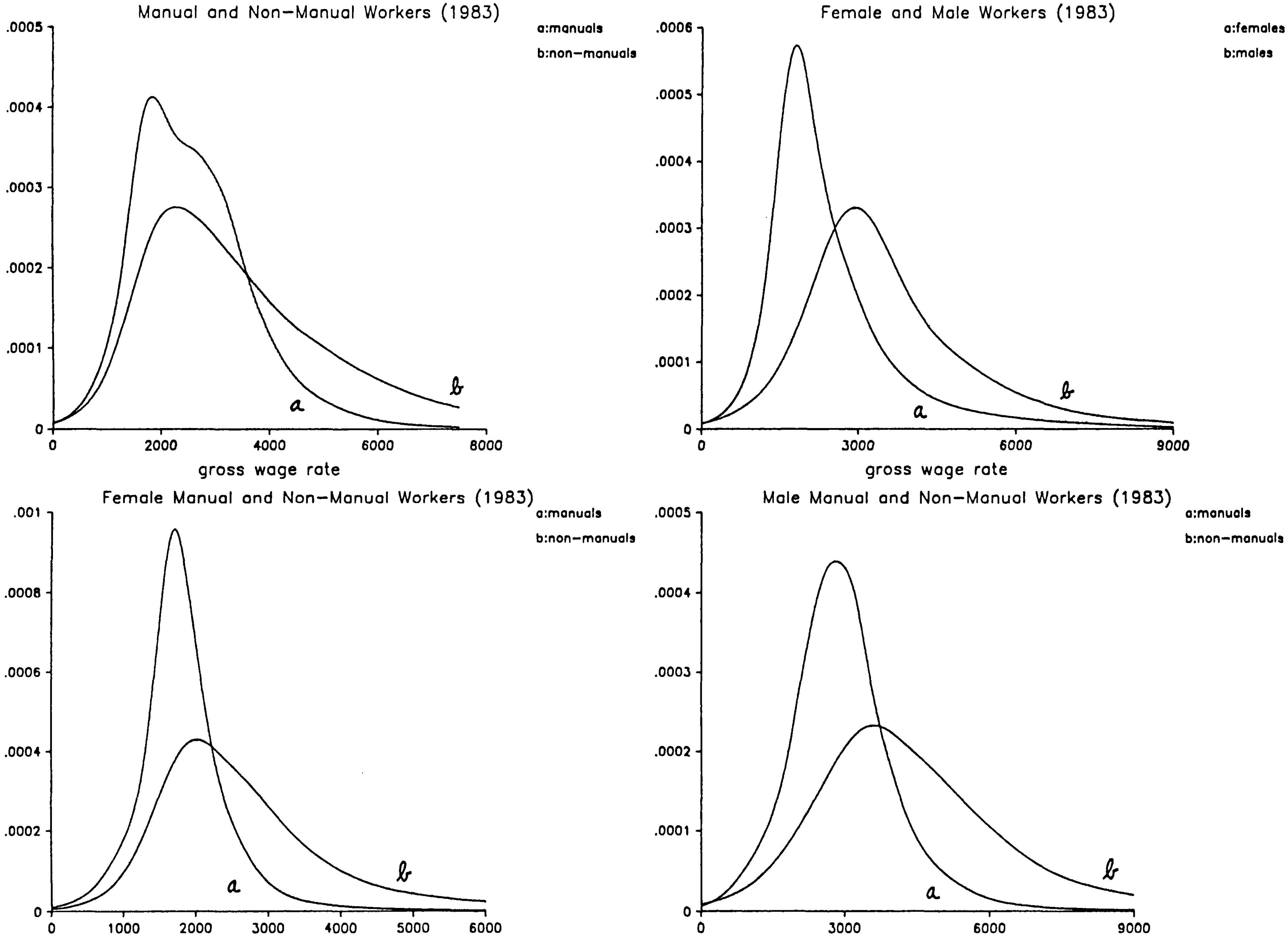


Figure 7a Adaptive kernel estimates

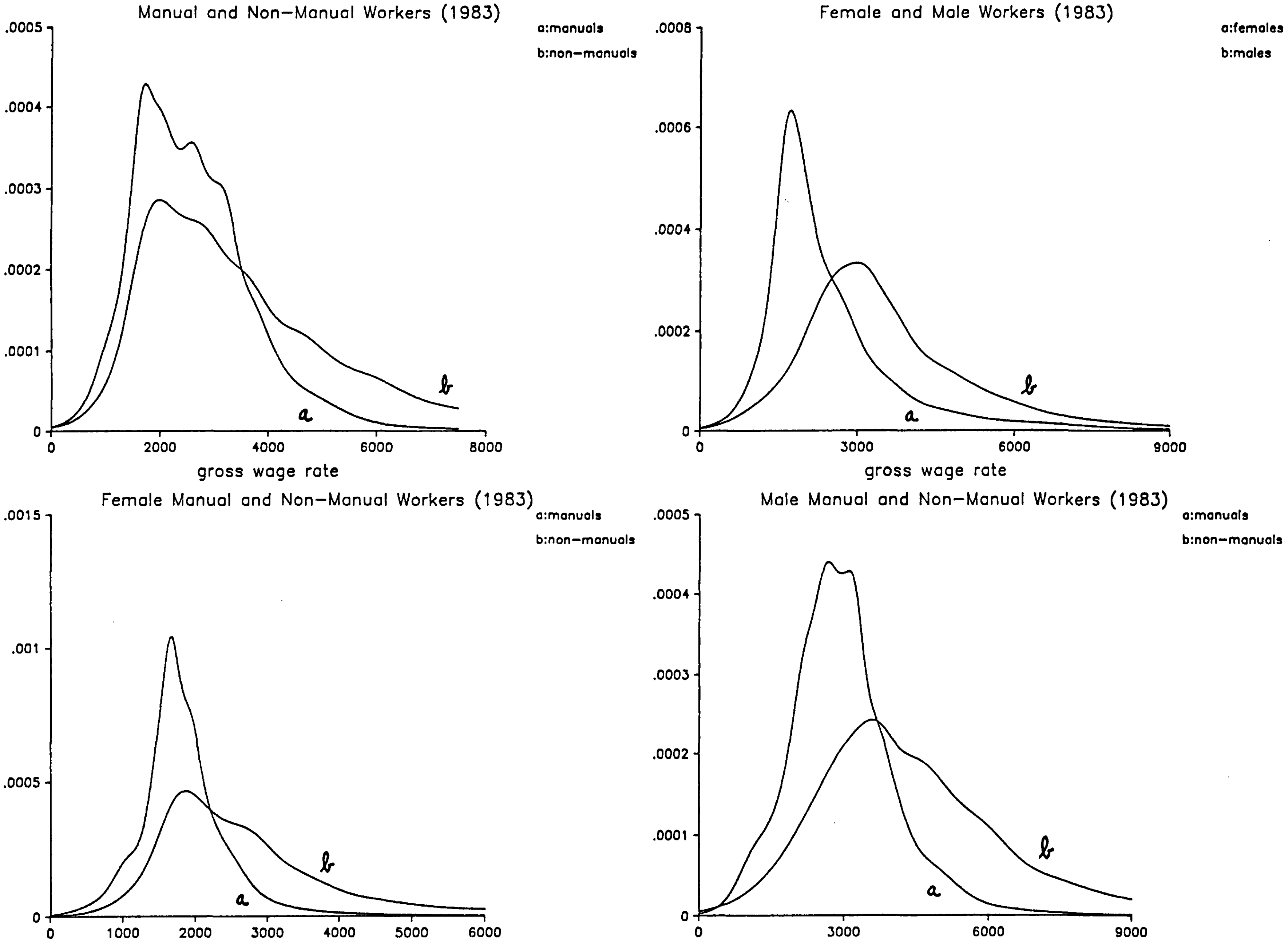


Figure 7b Adaptive kernel estimates; values of h decreased by around 35%

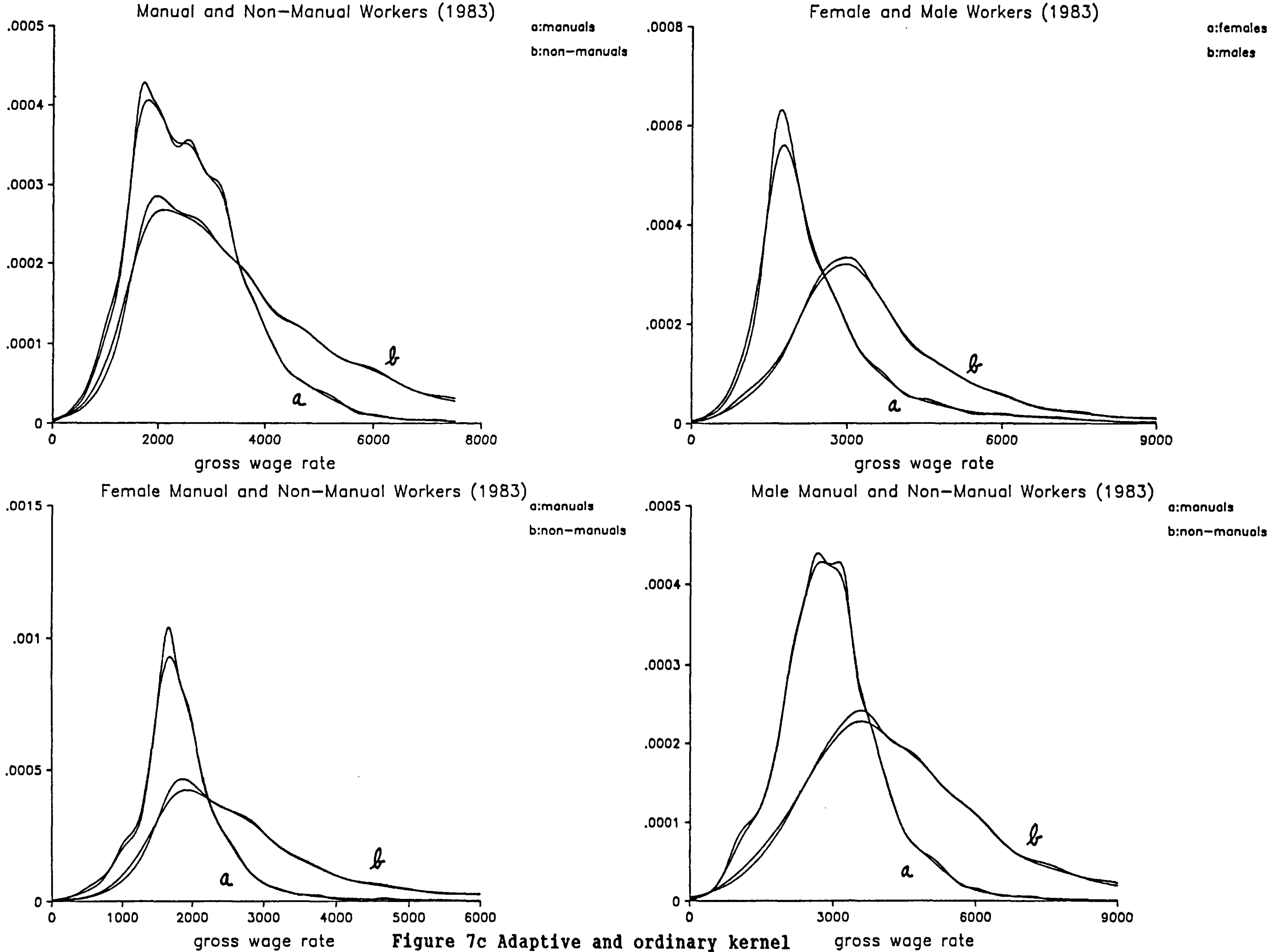


Figure 7c Adaptive and ordinary kernel estimates; values of h as in Figure 7b

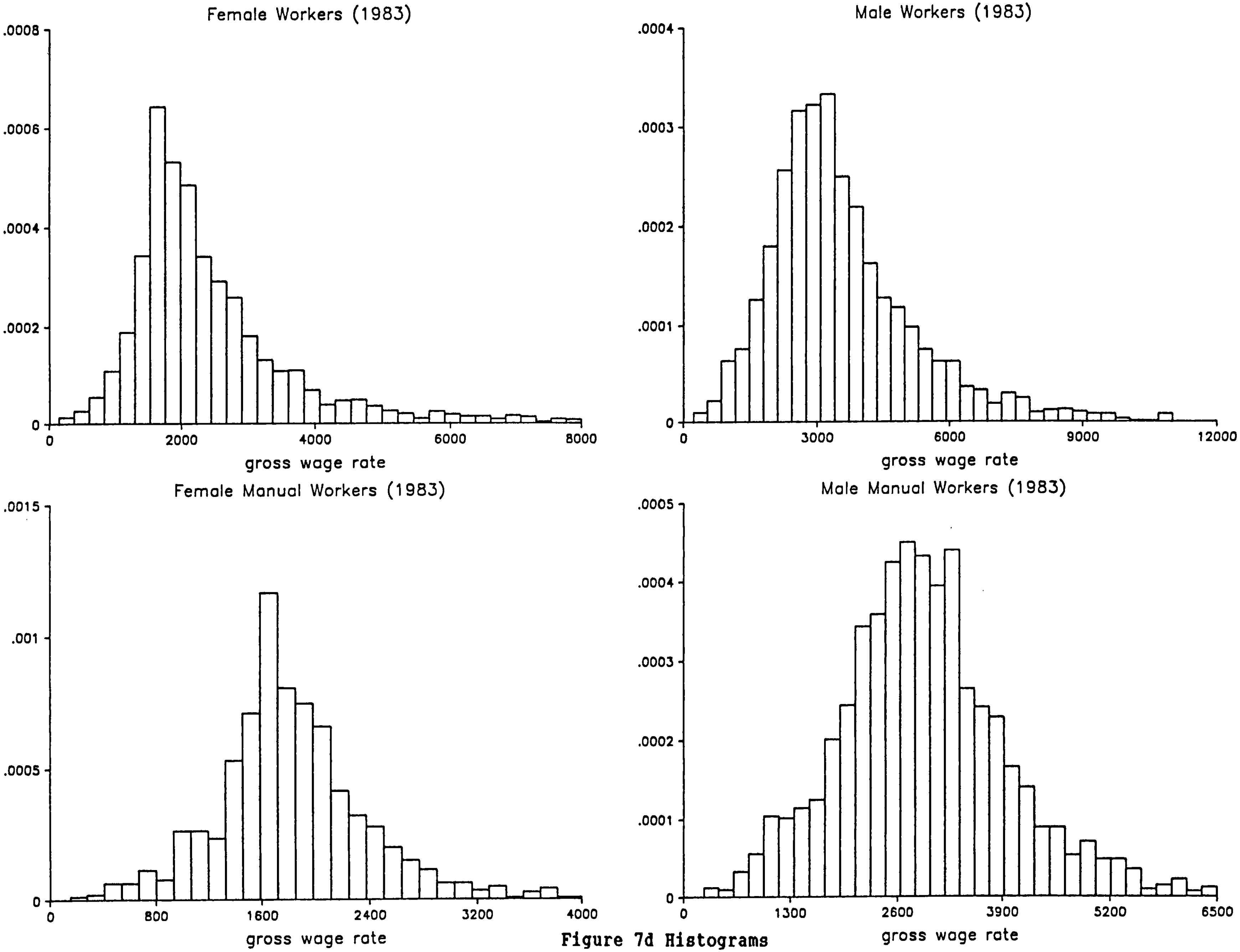


Figure 7d Histograms

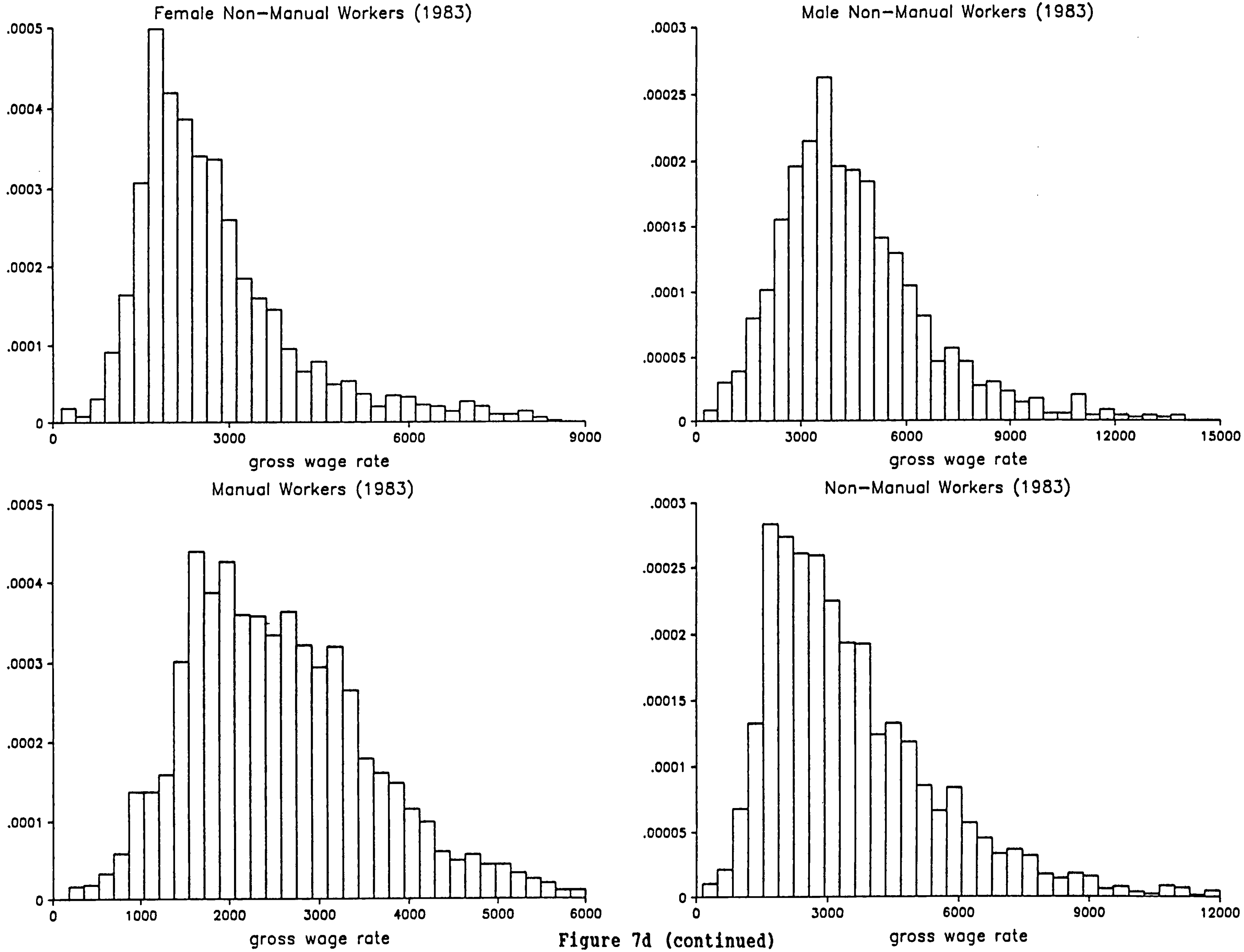


Figure 7d (continued)

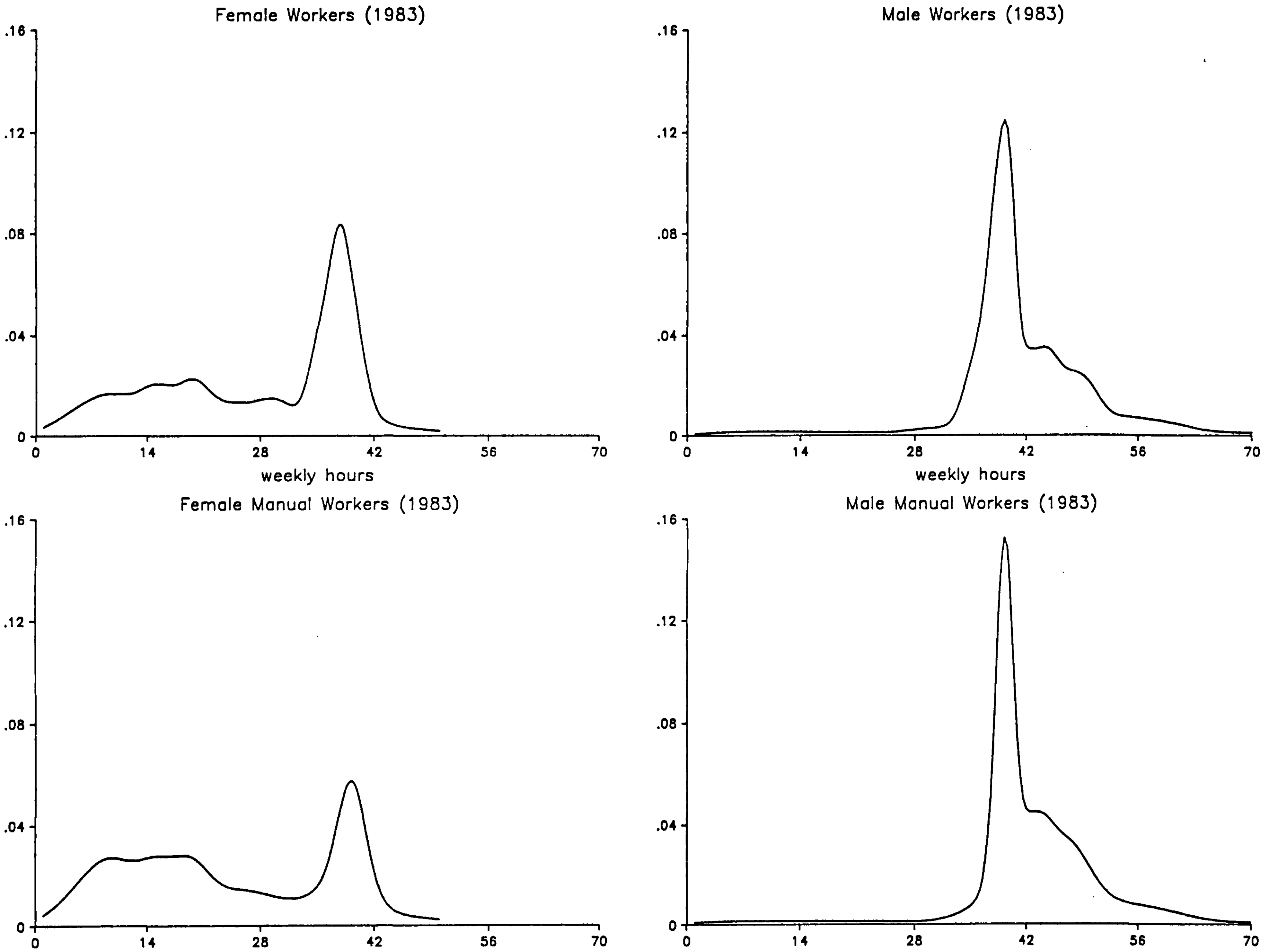


Figure 8a Adaptive kernel estimates

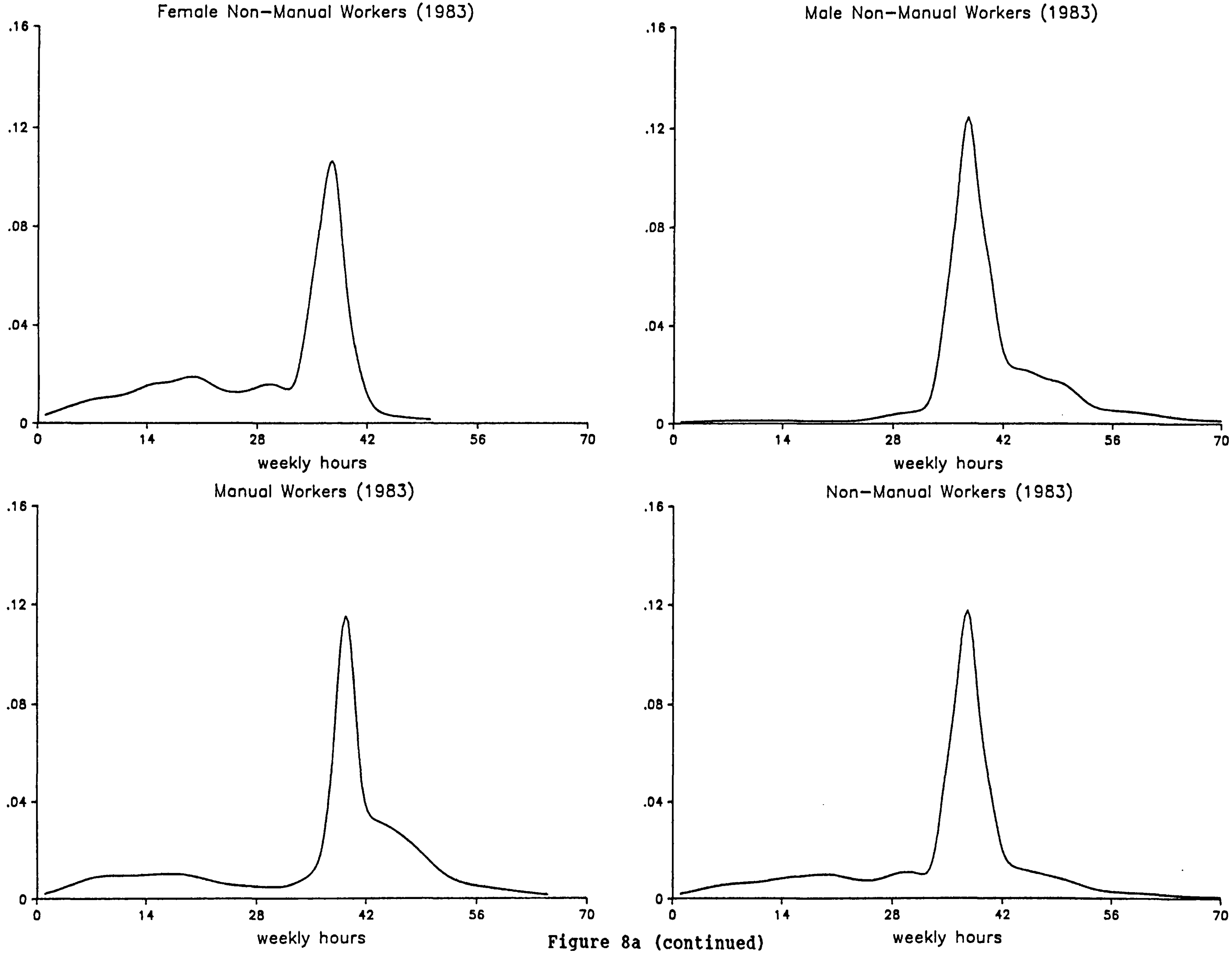


Figure 8a (continued)

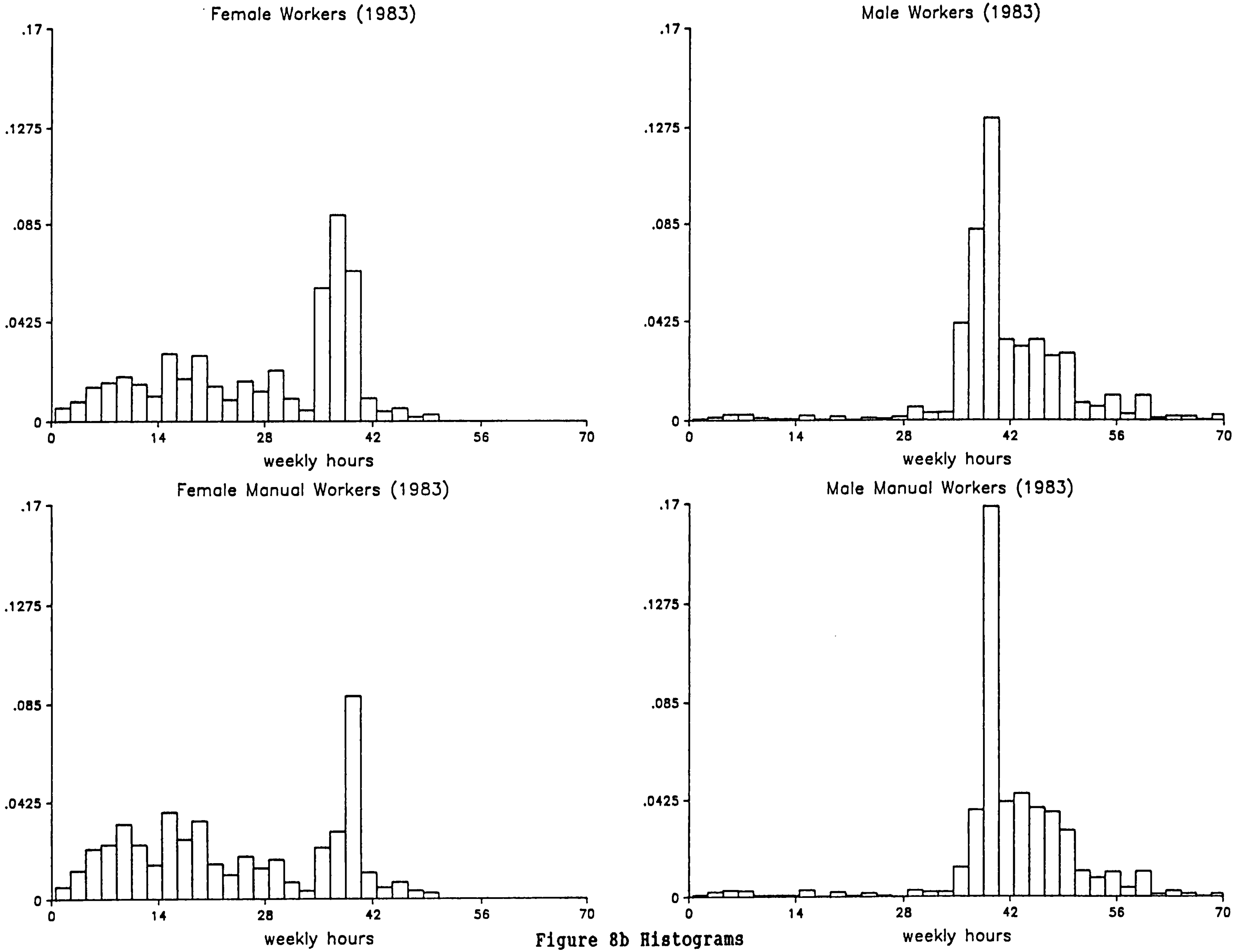


Figure 8b Histograms

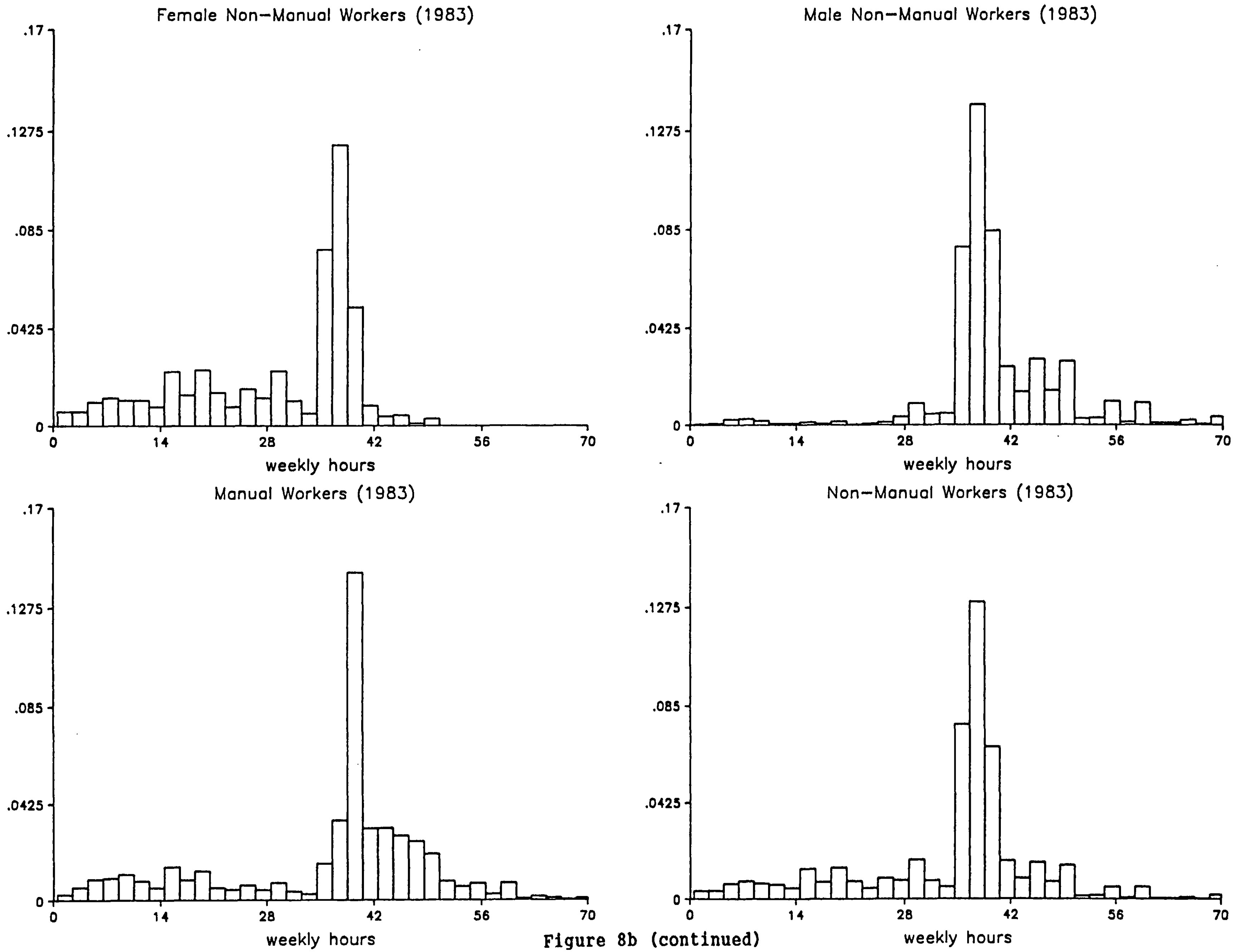


Figure 8b (continued)

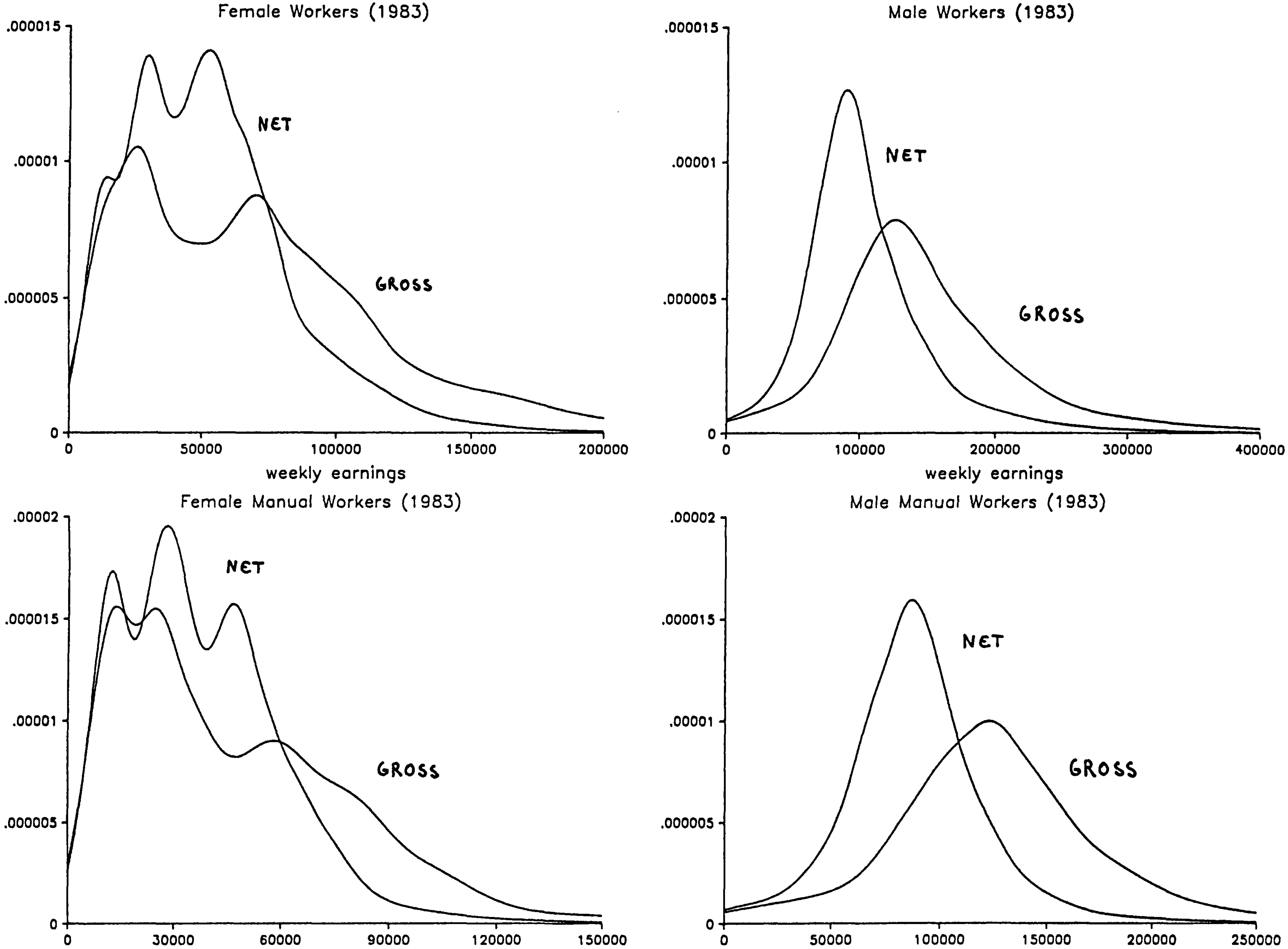


Figure 9a Adaptive kernel estimates; gross and net weekly earnings

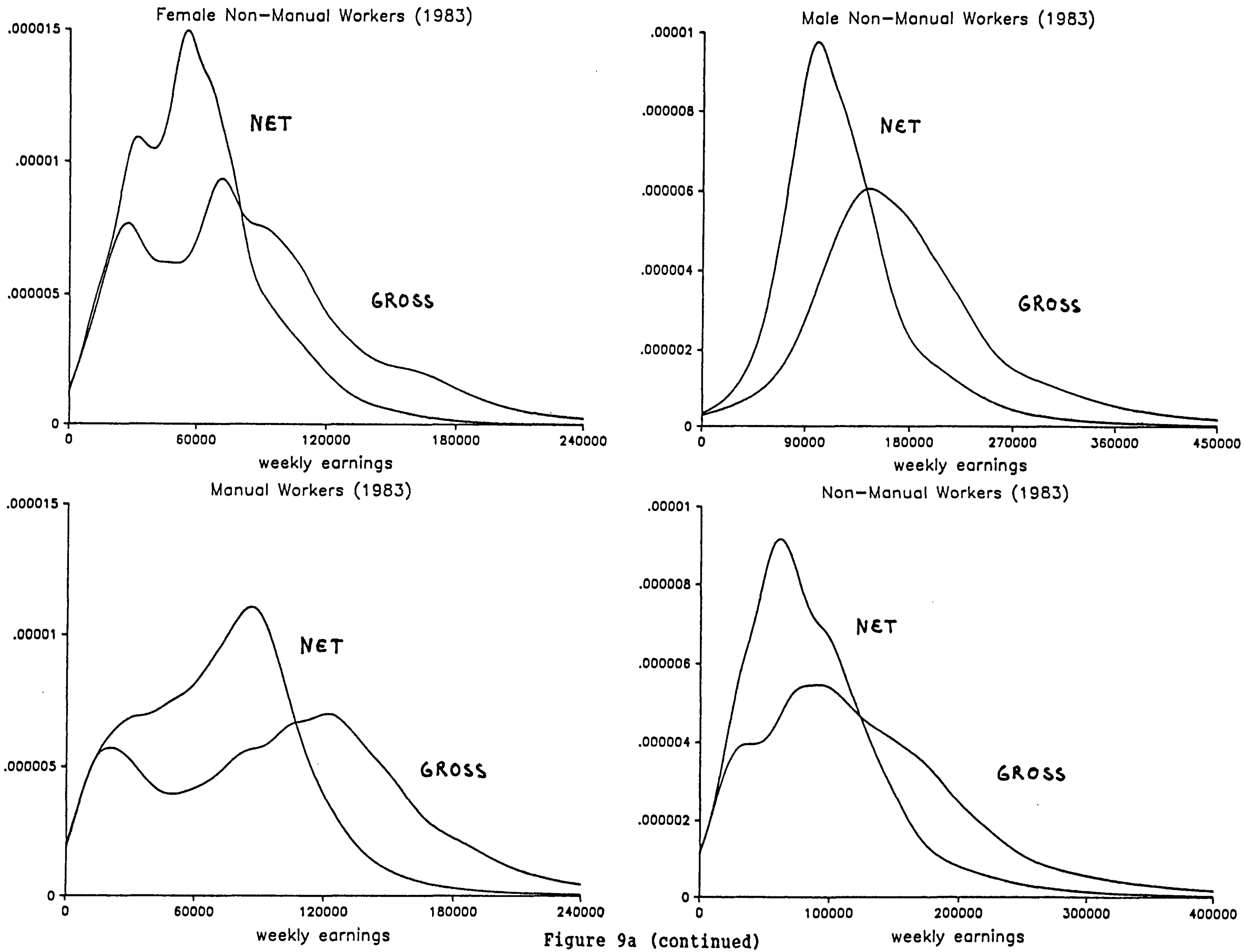


Figure 9a (continued)

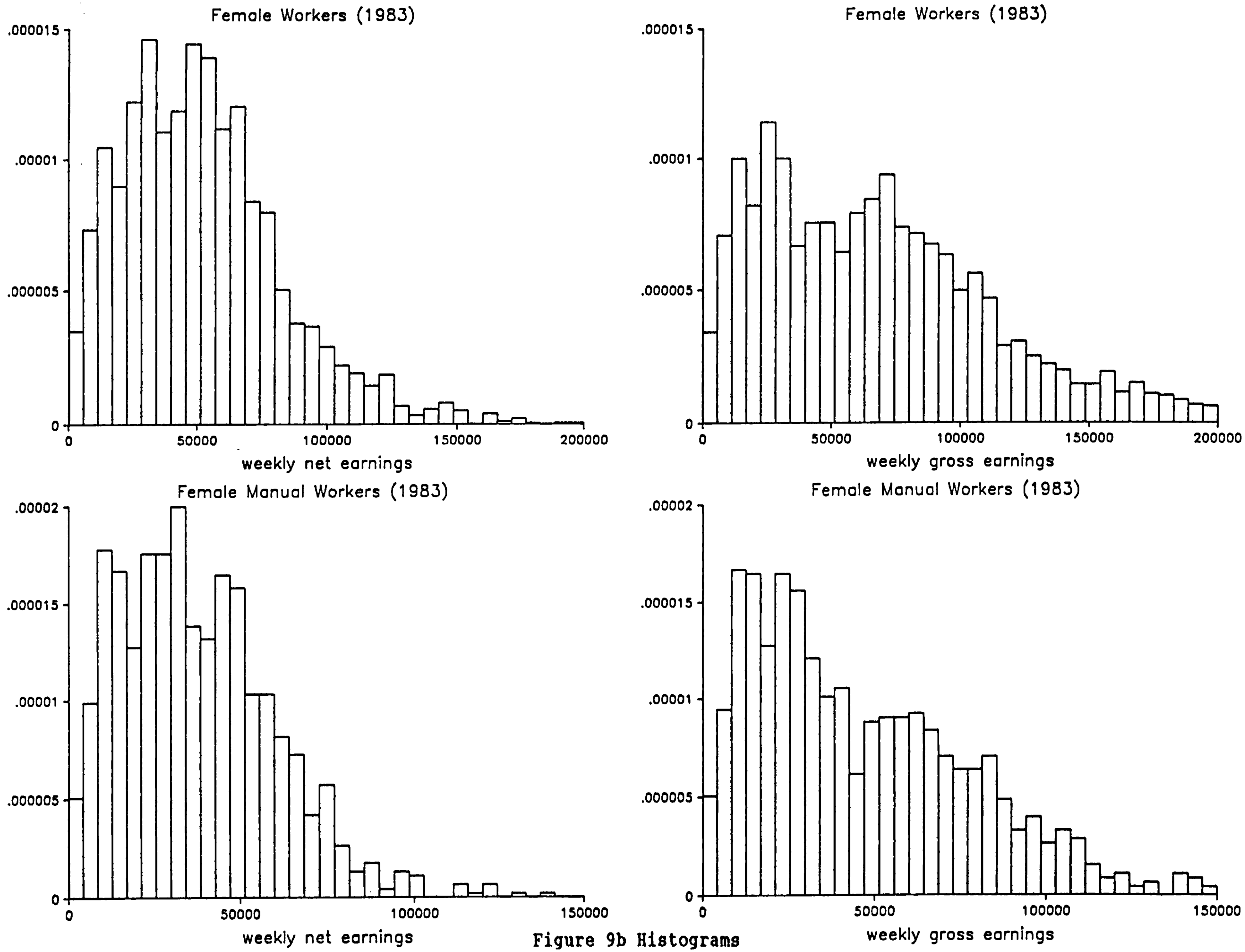


Figure 9b Histograms

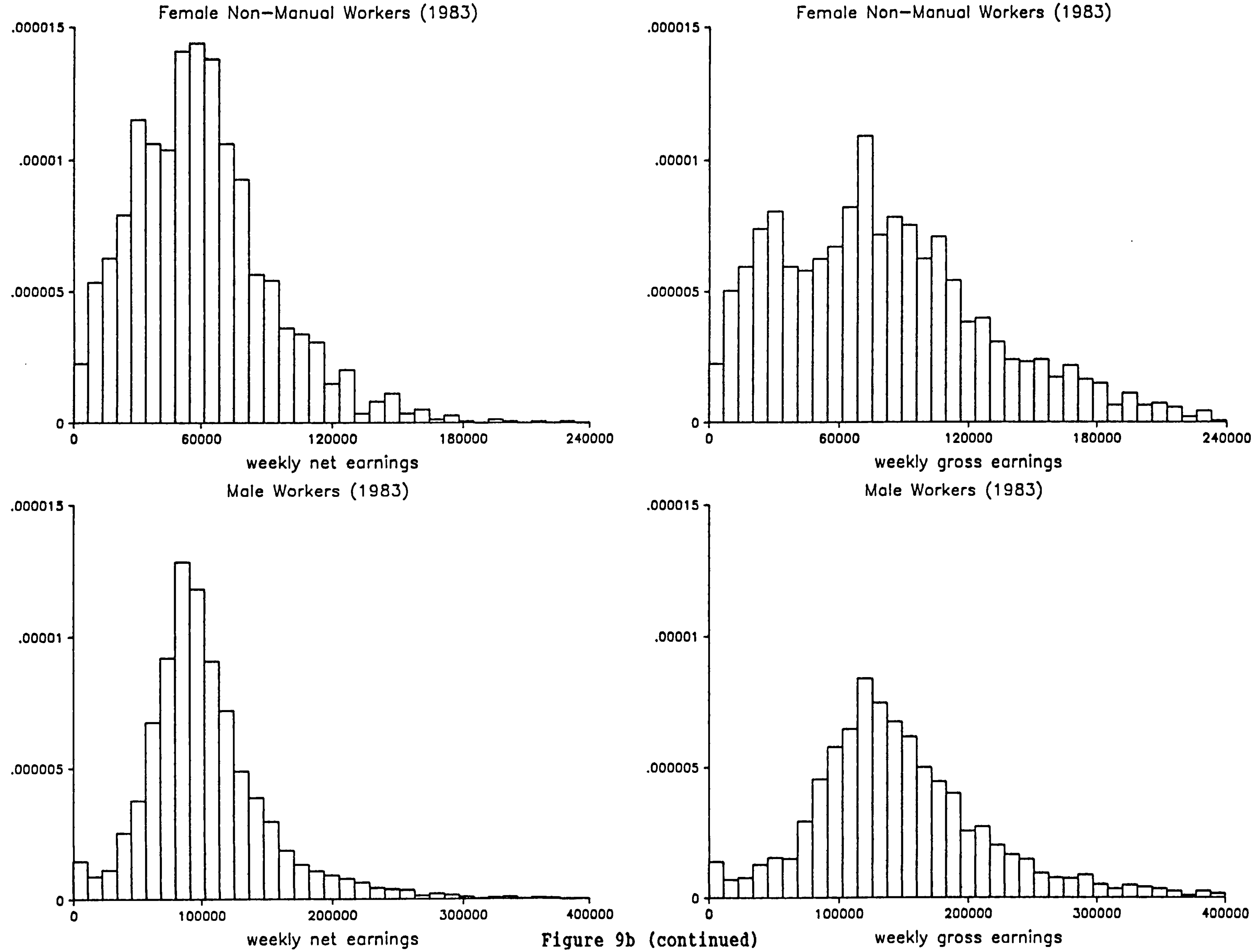


Figure 9b (continued)

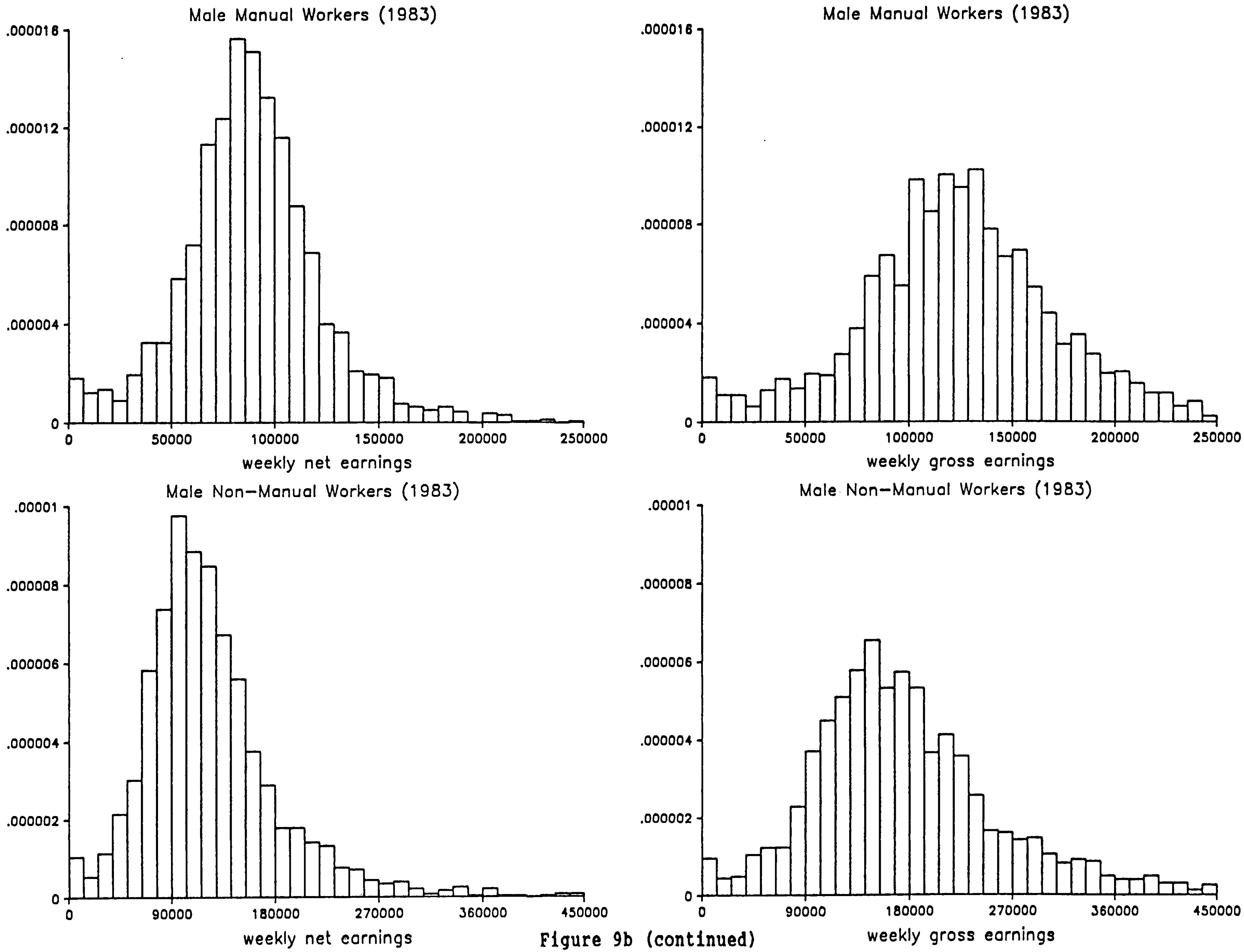


Figure 9b (continued)

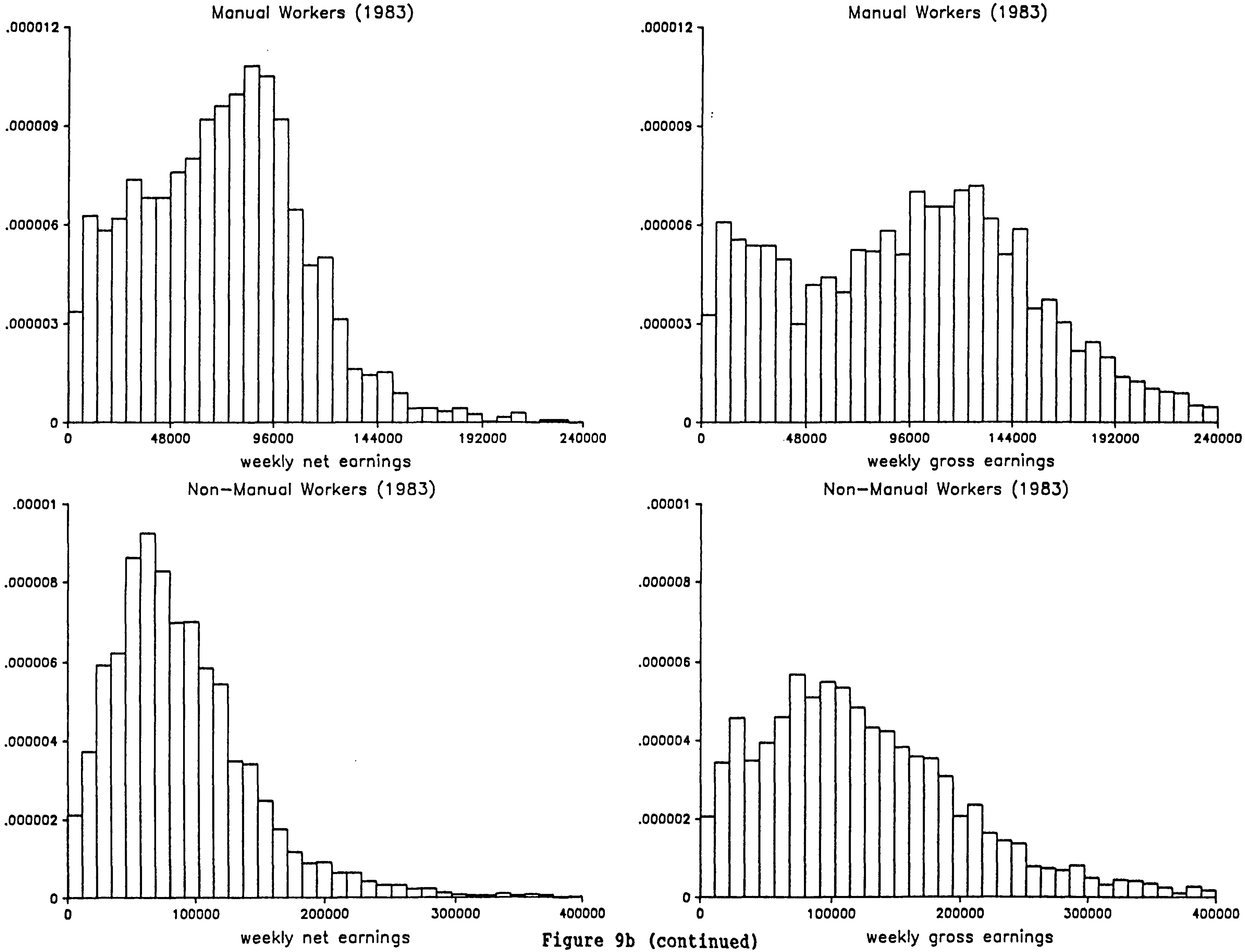


Figure 9b (continued)

5. Distributions over Time

So far we have been concerned with estimating distributions for a single year. Such estimates may be misleading for two reasons. Firstly, there may be deficiencies in the specific sample or the year to which the sample relates may be atypical. Secondly, we do not know whether the distributions under consideration are stable across the years, i.e., the shape of the density functions may change (either randomly or according to a pattern). One therefore has to be cautious in drawing conclusions from what has been done in the last two sections. However, the FES is a time-series of cross-sectional data, which allows us to study distributions over time.

In this section we will explore the evolution, over the period 1971-85, of the distributions of gross wage rates and of weekly hours of work. (We will comment briefly on the distribution of personal income below.) Before proceeding with the data analysis, observe first that the mean and the standard deviation of the gross wage rate data increase steadily from 1970 to 1985. The following figures are taken from Table 1a in the Appendix (μ and σ denote, respectively, the mean and the standard deviation, in £, of the data):

	1971	1973	1975	1977	1979	1981	1983	1985
μ	0.60	0.78	1.16	1.49	1.91	2.72	3.18	3.67
σ	0.37	0.52	0.62	0.77	1.03	1.82	1.98	2.55
σ/μ	0.62	0.67	0.54	0.52	0.54	0.67	0.62	0.69

Hence, in order to compare gross wage rate distributions for different years with each other, one has to normalise the data. In principle, one could deflate the earnings data in the annual samples by a price index and focus on the evolution of real gross wage rates. In this case the value of the density estimates would, however, depend on whether we have chosen the price indices properly. Consumption behaviour typically varies across groups of households, so that workers having in a particular year the same labour income do not necessarily also have the same purchasing power with respect to the price system in a base year. One therefore should not deflate all data in a sample by the same price index. The choice of the price indices requires much care; mistakes made at this stage may introduce substantial bias into the estimates.

To avoid the above difficulties, we made the estimates independent of the scale of the data by simply dividing all wage rates in a sample by its arithmetic mean. Let $\bar{x}$ denote the arithmetic mean of n real observations $x_1,\ldots,x_n$ drawn from an unknown distribution with density ρ; let μ and σ denote, respectively, the mean and the standard deviation of ρ. Then a density estimate for the data points $x_1/\bar{x},\ldots,x_n/\bar{x}$ is an estimate of the density $\mu\cdot\rho(\mu\cdot x)$, whose mean and standard deviation are given by 1 and σ/μ, respectively. In the following we will use the term "normalised" as an abbreviation for "normalised with respect to the mean".

We remark that the distribution of household net income in the FES was studied by K. Hildenbrand and W. Hildenbrand (1986). Using the "discrete maximum penalised likelihood" method (see Subsection 3.2), the authors estimated the net income density on the sample of "all households" for the years from 1969 until 1981. The estimations show that the income distribution is bimodal. Furthermore, it turned out that the left peak increased while the right peak (with two exceptions) decreased over the period 1969-81. The time series of distributions begins with densities whose second maximum dominates the first and ends with densities where the relation between the two maxima is exactly the reverse (see also Figure 5 in W. Hildenbrand, 1989a). In other words, the percentage of households with low income (relative to the mean) steadily increased during the 1970s.

Unfortunately, the authors do not say much about this change in the income distribution. They only remark: "the Lorenz curves (and hence the Gini coefficient) of the empirical data do not differ substantially for the years 1969 to 1981. A parametric estimation (e.g., lognormal) of the normalized data in every year leads approximately to the same density" (p. 256).[8] It is open to discussion whether or not the change in the distribution was "substantial". For instance, it appears that the proportion of families with an income below the Supplementary Benefit level increased by around a quarter between 1979 and 1981 (Department of Health and Social Security, 1983).

It is interesting to ask whether there was a similar change in the distribution of gross wage rates during the 1970s. The precise formulation of the question which we want to explore in this section is as follows: Let ρ_t and μ_t denote, respectively, the gross wage rate density and the mean of this density in year t, where t = 1971,...,1985. Does there then exist a density ρ with mean equal to 1, so that (at least approximately) $\mu_t \rho_t(\mu_t x) = \rho(x)$ for all x and t?

Observe from the above figures that the standard deviations of the normalised gross wage rate data, i.e., the coefficients of variation σ/μ, are fairly stable; the mean and the standard deviation of the eight numbers are 0.61 and 0.06, respectively. But notice also that the coefficients of variation do not change randomly between 1970 and 1985: they are lower in the period from 1975 to 1980 than in the remaining years (see also Table 1a in the Appendix).

In the case of labour supply we do not have to normalise the data. We want to investigate whether or not the distribution of labour supply did change in the time period under consideration. A brief glance at Table 1a in the Appendix shows that the means and standard deviations of the labour supply data are very stable. The arithmetic mean of the sample means for the years from 1971 to 1985 is 36.50; the standard deviation of the 15 numbers is 0.71. Thus, the sample means deviate on average by only 2 per cent from its arithmetic mean. The sample standard deviations spread slightly more around its mean: the arithmetic mean and the standard deviation of the 15 numbers are 12.75 and 0.34, respectively, giving us an average deviation from the mean of 2.7 per cent. But notice that just as the coefficients of variation of the gross wage rate data, the values of the two statistics do not change randomly over time, which conflicts with the hypothesis that the labour supply distribution did not change.

The sample means decrease in the period from 1971 to 1975, appear to change "randomly" between 1975 and 1979, and decrease during the years 1979-84 again. Comparison of the figure for 1971 with that for 1984 shows a decrease in mean labour supply of 5.8 per cent (we will return to this observation). If we approximate the time series of sample standard deviations by a smooth curve, then we obtain a curve which is strictly increasing between 1971 and 1974, flat in the period from 1974 until 1981, where the sample standard deviations appear to change randomly, and from 1981 onwards again strictly increasing. We observe between 1971 and 1985 an increase in the spread of the data around its mean of 11.4 per cent. Looking at each second year, the following picture emerges (μ and σ denote, respectively, the mean and the standard deviation of the labour supply data):

	1971	1973	1975	1977	1979	1981	1983	1985
μ	37.62	37.20	36.99	36.66	36.81	35.82	35.57	35.78
σ	12.05	12.49	12.62	12.76	13.02	12.64	13.08	13.42
σ/μ	0.320	0.336	0.341	0.348	0.354	0.353	0.368	0.375

Let us now turn to the estimations. We begin again with an overview of the diagrams which are displayed on pages 133-144. The distribution of gross wage rates (resp. hours of work) was estimated on the sample of "all workers" for each odd numbered year from 1971 to 1985. Figure 10 shows adaptive kernel estimates of the distributions of nominal gross wage rates. The diagram confirms what we have already said at the beginning of this section: if we want to compare wage rate densities for different years with each other, we have to normalise the data. In Figure 11 the empirical cumulative distribution functions of the normalised gross wage rates (i.e., mean gross wage rate equal to one) are plotted. Adaptive kernel estimates and histograms for the normalised data are drawn in Figure 12. Finally, we see in Figure 13 adaptive and ordinary kernel estimates of the labour supply distributions. As always, the ordinary kernel estimates served as the pilot estimates required to obtain the adaptive kernel smoothers.

Notice that the labour supply density obtained by the adaptive kernel method has a much higher peak than that obtained by the ordinary kernel method (especially in the years 1971-77). A brief reminder why this is the case may be useful: The adaptive kernel estimator is defined in such a way that (i) the *local window width* h_i corresponding to the observation x_i is a strictly decreasing function of $\hat{\rho}(x_i)$, where $\hat{\rho}$ denotes the ordinary kernel estimate of the labour supply density, and (ii) the geometric mean of the h_i is equal to the *global window width* h which is used to compute the pilot estimate (h=1.6 in Figure 13). Hence the local window width h_i at the maximum of the ordinary kernel estimate is smaller than h. This, in turn, implies that the adaptive kernel method gives more weight to observations near the maximum than the ordinary kernel method.

It is interesting to observe how remarkably stable the labour supply distributions are in the years 1971-85. In the first half of this period the estimated densities differ only slightly in its lower and upper range. Figure 13d shows that the density estimates for the years 1971, 1973, 1975 and 1977 are so similar that it appears as if one diagram was copied four times. However, the tails of the density functions do not change randomly over the years. A closer look at Figure 13 reveals that the proportion of individuals working less than 25 hours per week increased during the 1970s.

Nevertheless, we think it is reasonable to conclude from the estimations that the data support the hypothesis of an "almost" constant labour supply distribution in the years 1971-77. The picture changes during the years from 1977 to 1985. We now observe a clear pattern of shifts: the peaks of the densities decrease and shift to the origin. Of course, this was to be expected from what has been said above about the sample means and standard deviations. In our opinion, however, these shifts are not really significant. At the very least, Figure 13 suggests that the distribution of labour supply did not change substantially between 1971 and 1985. Table 1 displays the evolution of some sample percentiles.

Table 1 Labour Supply Distributions

Sample percentiles: all workers

YEAR	1%	5%	10%	25%	50%	75%	90%	95%	99%
1971	5	10	20	35	40	44	50	55	65
1973	4	10	18	35	40	43	50	55	65
1975	5	10	18	35	40	42	50	55	68
1977	4	10	16	35	40	42	50	55	68
1979	4	10	16	35	40	43	50	55	70
1981	4	10	16	32	38	40	48	54	67
1983	3	9	15	32	38	41	48	54	69
1985	3	8	15	32	38	42	50	55	70

The distributions of the normalised gross wage rates are considerably less stable. Figures 11 and 12 show that the distributions differ essentially in the range of normalised gross wage rates which are not greater than 1.6 (approximately 90 per cent of the observations are contained in this range). In the following we will use the term "wage rate" as an abbreviation for "normalised gross wage rate".

As we see in the first diagram of Figure 11a, the empirical distribution functions for 1971 and 1973 are almost identical for wage rates not greater than 1.05; in this range of wage rates the two distribution functions also lie above the distribution functions for 1975 and 1977. For wage rates greater than 1.05 and smaller than 1.7 the cdf for 1973 lies above that for 1971. Over the interval [1.5,2] the graphs of the two functions approach each other and for wage rates greater than 2 they are indistinguishable by eye.

For wage rates greater than 1.7 the graphs of the distribution functions for 1975 and 1977 are also almost equal; it is, however, visible that they lie below the curves for 1971 and 1973. For wage rates smaller than 0.75 the 1977 cdf lies below the 1975 cdf; over the interval [0.75,0.90] the two distribution functions do not differ, and for wage rates greater than 0.90 the former function assumes larger values than the latter. In the interval [1.15,1.40] the distribution functions for 1973 and 1975 are approximately equal; in the range of wage rates which are greater than 1.40 the graph of the former function lies below that of the latter. Thus, in the range of wage rates which are grater than 1.15 the distribution function for 1977 assumes the largest values and that for 1971 the smallest.

Summing up, during the years from 1971 to 1977 we observe in the population a decrease in the proportion of workers receiving low wage rates (relative to the mean) as well as a decrease in the proportion of individuals receiving high wage rates.

The empirical distribution functions of the normalised gross wage rate data for the years 1979, 1981, 1983 and 1985 are plotted in the second diagram of Figure 11a. Again the distribution functions change not randomly over time but according to a pattern. In the range of wage rates which are smaller than 1.1 the distribution function for 1979 (resp. 1985) lies below (resp. above) the other three functions; the distribution functions for 1981 and 1983 do not differ very much, but the former function assumes in general smaller values. Over the interval [1.10,1.35] the graphs of the four functions are almost identical. In the interval [1.35,2.00] the distribution function for 1979 (resp. 1985) lies above (resp. below) the other functions. The cdf for 1983 differs only very slightly from that for 1985 and lies below that for 1981. For wage rates greater than 2 the graphs of the distribution functions for 1981, 1983 and 1985 are almost indistinguishable by eye, while it is still visible that the cdf for 1979 assumes larger values than the other three functions.

Hence, over the period from 1979 to 1985 we are faced with an increase in the proportion of workers earning low wages and an increase in the proportion of individuals earning high wages (relative to the mean).

The density estimates in Figure 12a show us the same changes as the empirical distribution functions. Roughly speaking, the densities in the first diagram of this figure shift to the right, while the densities in the second diagram shift to the left. (It should not be forgotten that the curves in Figures 11 and 12 shift in such a manner that its mean does not change.) Notice that the wage rate distributions shift between 1977 and 1985 in the same direction as the labour supply distributions in this period. Observe also that we are faced with the following pattern of shifts if we exclude the density estimates for the years 1971 and 1973 from the eight curves plotted in Figure 12b: the modes of the six remaining densities move to the left.

These changes are also inferable from the sample percentiles given in Table 2 on the next page. Reading down the first five columns of the table, we see that the first, fifth, tenth, twenty-fifth and fiftieth percentiles first increase and then decrease. Looking down the last three columns, we see that the ninetieth, ninety-fifth and ninety-ninth percecntiles first decrease and then increase.

Table 2 Distribution of Normalised Gross Wage Rates
Sample percentiles: all workers

YEAR	1%	5%	10%	25%	50%	75%	90%	95%	99%
1971	0.242	0.361	0.440	0.608	0.874	1.220	1.650	2.039	3.259
1973	0.247	0.370	0.451	0.613	0.874	1.203	1.629	2.032	3.171
1975	0.259	0.397	0.479	0.657	0.897	1.207	1.588	1.956	2.899
1977	0.268	0.419	0.520	0.676	0.897	1.188	1.566	1.935	2.823
1979	0.262	0.423	0.520	0.654	0.897	1.208	1.598	1.915	2.781
1981	0.238	0.387	0.479	0.620	0.868	1.210	1.651	2.050	3.000
1983	0.226	0.377	0.471	0.612	0.864	1.196	1.686	2.079	3.171
1985	0.216	0.382	0.463	0.597	0.843	1.206	1.708	2.096	3.265

If one computes the ratio of the wage rate at the top decile to that at the bottom decile, then one obtains the following time-series for the so-called *decile ratio*:

1971	1973	1975	1977	1979	1981	1983	1985
3.750	3.612	3.315	3.017	3.073	3.447	3.580	3.689

As we see, the decile ratio is decreasing in the years 1971-77 and increasing in the years 1977-85. In other words, the *inequality* in the distribution of gross wage rates is first decreasing and then increasing.

It is interesting to remark that Atkinson and Micklewright (1992, pp. 86-87) make the same observation using the data of the British New Earnings Survey (NES). Recall that we mentioned the NES already in Section 2. Both the FES and the NES are time-series of cross-sectional data. The NES, however, is much larger than the FES: the size of the annual NES sample is around 170,000 persons while the FES has an annual sample of around 7,000 households. Furthermore, the NES obtains its information from employers. One may therefore take the view that the NES data are more reliable than those of the FES where interviewers visit the households. But it appears that this is not the case (we refer the reader to our brief discussion of the FES data in Section 2).

Atkinson and Micklewright consider the variable "gross weekly earnings" and compute the decile ratio in the NES samples for the years 1968-89 (the NES data were first collected in 1968). It turns out that the decile ratio fell (resp. rose) in the years 1968-77 (resp. 1977-89) by 21.2 (resp. 17.2) per cent, namely from 3.68 in 1968 to 2.9 in 1977 (resp. from 2.9 in 1977 to 3.4 in 1989); Figure 4.1 in Atkinson and Micklewright (1992, p. 86) shows that the decile ratio for 1971 is 3.4, approximately, and that for 1985 is 3.2, approximately.

The authors provide an explanation why the decile ratio decreased in the period 1968-77. But before quoting their explanation, let us first have a look at the data underlying our estimates.

The next tables (pp. 127-129) shed some light on the factors which possibly caused the shifts in the distributions. Table 3 presents sample proportions; Tables 4-8 contain sample sizes and sample ratios. Recall from the preceding section that an individual is classified as a *full-time* worker if his or her *normal hours* are greater than 30.

As we see from Tables 3 and 4, the proportion of females in the labour force increased continuously over the period 1971-85 (from around 39 to 45 per cent). The sample proportion of part-time female (resp. male) workers increases (resp. decreases) during the years 1971-77 by 3.2 (resp. 0.5) percentage points; over the period 1977-85 the sample proportion of part-time female (resp. male) workers increases by 1.4 (resp. 1.0) percentage points. Loosely speaking, the proportion of part-time males in the labour force is of the order of 3 per cent. Thus, the increase in part-time work as shown in Table 5 is essentially attributable to a rise in part-time female labour supply. (If we divide the number of part-time male workers by the number of full-time workers in each year, this quotient decreases between 1971 and 1977 from 0.036 to 0.031, and increases during the years 1979-85 from 0.036 to 0.044.)

In Tables 6 and 7 we can see that the sample ratio of females to males increases more for non-manual than for manual workers. However, the change in the sample proportion resulting from a change in the sample ratio is a decreasing function of the sample ratio.

The proportion of females among both types of workers increased over the period 1971-85 by around 4 percentage points, namely from 0.509 to 0.548 in the population of "non-manual workers" and from 0.299 to 0.338 in that of "manual workers". Notice also that the proportion of females in the subsamples of full-time manual workers decreases in the years 1971, 1973 and 1975 (resp. 1979, 1981 and 1983), while we observe for full-time non-manual workers between 1971 and 1977 an increase in the sample ratio of females to males and a subsequent decrease until 1983. Moreover, a brief glance at the sample sizes in Tables 6 and 7 shows a general switch from manual to non-manual occupations.

The proportion of non-manual workers in the labour force grew rapidly in these years, namely from 0.326 to 0.420 in the population of "male workers" and from 0.541 to 0.633 in that of "female workers". Thus, we can conclude that the overall rise in female labour supply was largely due to females entering the labour market as non-manuals (many of them working part-time). Finally, Table 8 shows that the sample ratio of non-manuals to manuals increases more between 1979 and 1985 than in the years from 1971 to 1977. The sample proportion of non-manual workers rises during the years 1971-77 (resp. 1979-85) from 0.409 to 0.442 (resp. from 0.462 to 0.515).

Table 3 Proportions of Subgroups in the Annual Samples of "All Workers"

	1971	1973	1975	1977	1979	1981	1983	1985
Part-time:								
Females	0.165	0.185	0.194	0.197	0.199	0.210	0.213	0.211
Males	0.029	0.027	0.027	0.024	0.028	0.030	0.031	0.034
Full-time:								
Females	0.220	0.215	0.218	0.224	0.227	0.226	0.225	0.235
Males	0.586	0.573	0.561	0.554	0.546	0.533	0.531	0.520

Table 4a Female and Male Workers Sample Sizes and Sample Ratios

YEAR	MALES	FEMALES	RATIO
1971	5109	3198	0.626
1973	4946	3298	0.667
1975	4906	3428	0.699
1977	4737	3447	0.728
1979	4379	3259	0.744
1981	4565	3538	0.775
1983	3837	2996	0.781
1985	3842	3100	0.807

Table 4b Full-Time Female and Male Workers[1)]

YEAR	MALES	FEMALES	RATIO
1971	4867	1829	0.376
1973	4724	1771	0.375
1975	4679	1813	0.387
1977	4537	1834	0.404
1979	4168	1736	0.417
1981	4322	1834	0.424
1983	3625	1540	0.425
1985	3609	1631	0.452

Table 5 Full-Time and Part-Time Workers[1)]

YEAR	FULL-TIME	PART-TIME	RATIO
1971	6691	1616	0.242
1973	6495	1749	0.269
1975	6492	1842	0.284
1977	6371	1813	0.285
1979	5901	1737	0.294
1981	6156	1947	0.316
1983	5165	1668	0.323
1985	5240	1707	0.326

Table 6a Non-Manual Female and Male Workers[1)]

YEAR	MALES	FEMALES	RATIO
1971	1668	1730	1.037
1973	1702	1777	1.044
1975	1745	1991	1.141
1977	1666	1951	1.171
1979	1632	1895	1.161
1981	1821	2146	1.178
1983	1698	1934	1.139
1985	1615	1961	1.214

Table 6b Full-Time Non-Manual Female and Male Workers[1)]

YEAR	MALES	FEMALES	RATIO
1971	1584	1125	0.710
1973	1619	1135	0.701
1975	1642	1224	0.745
1977	1577	1217	0.772
1979	1536	1174	0.764
1981	1711	1289	0.753
1983	1592	1140	0.716
1985	1508	1176	0.780

1) Sample sizes and sample ratios.

Table 7a Manual Female and Male Workers[1)]

YEAR	MALES	FEMALES	RATIO
1971	3441	1468	0.427
1973	3244	1521	0.469
1975	3161	1437	0.455
1977	3071	1496	0.487
1979	2747	1364	0.497
1981	2744	1392	0.507
1983	2139	1062	0.496
1985	2232	1139	0.510

Table 7b Full-Time Manual Female and Male Workers[1)]

YEAR	MALES	FEMALES	RATIO
1971	3283	704	0.214
1973	3105	636	0.205
1975	3037	589	0.194
1977	2960	617	0.208
1979	2632	562	0.214
1981	2611	545	0.209
1983	2033	400	0.197
1985	2101	455	0.217

Table 8a Manual and Non-Manual Workers[1)]

YEAR	MANUALS	NON-MANUALS	RATIO
1971	4909	3398	0.692
1973	4765	3479	0.730
1975	4598	3736	0.813
1977	4567	3617	0.792
1979	4111	3527	0.858
1981	4136	3967	0.959
1983	3201	3632	1.135
1985	3371	3576	1.061

Table 8b Full-Time Manual and Non-Manual Workers[1)]

YEAR	MANUALS	NON-MANUALS	RATIO
1971	3987	2709	0.679
1973	3741	2754	0.736
1975	3626	2866	0.790
1977	3577	2794	0.781
1979	3194	2710	0.848
1981	3156	3000	0.951
1983	2433	2732	1.123
1985	2556	2684	1.050

1) Sample sizes and sample ratios.

These two changes in the composition of the annual samples, i.e., the increase in female labour supply and the switch from manual to non-manual occupations, explain the shifts in the densities of the labour supply distributions. [Recall from Section 4 that females (resp. non-manuals) work on average less hours per week than males (resp. manuals); see also the tables in the Appendix.]

The factors which caused the shifts in the wage rate densities are not that easily to detect. An increase in the sample proportion of workers earning low wages will shift the empirical density somewhat nearer to the origin; and as the diagrams of Figure 7 and the summary statistics in the Appendix show, females receive lower gross wage rates than males. It seems therefore reasonable to conclude, that the observed shifts in the densities over the period 1979-85 are attributable to the high proportion of female workers in the FES samples from the end of the 1970s onwards.[9)] Our brief data analysis does not, however, explain the shifts in the distributions during the years 1971-77.

We remark that Atkinson and Micklewright (1992, pp. 86-87) provide the following explanation for the fall in the decile earnings ratio in Britain over the period 1968-77 (see pages 125-126): "this fall...was associated in part with the improvement of the relative earnings position of women (it was over this period that the Equal Pay legislation was implemented), but there was also a reduction in dispersion for male workers: the decile ratio for men aged twenty-one and over fell from 2.46 in 1970 to 2.32 in 1977...A role is likely to have been played by the post-1970 incomes policies which incorporated flat-rate, rather than percentage, elements...such as the 1975 policy of increases of £6 a week for all workers, except those earning more than £8,500 a year (within the top percentile)."

Let us suggest at the end of the section possible directions for future research in this area. Firstly, one could explore in a subsequent study the inequality in the gross wage rate distributions. It would be desirable to obtain a (partial) ranking of the distributions within an analysis that pays attention to both inequality between workers and the extent to which greater inequality can be compensated by higher average real wage rates.[10)] Secondly, it would be interesting to examine the lower range of the gross wage rate distributions in more detail. This can be done by applying well-known poverty measures to the estimates. It should be explored whether poverty within the work force, measured by low hourly wages, actually decreased (resp. increased) during the years 1971-77 (resp. 1977-1985).[11] (Although the observed shifts in the wage rate distributions may be interpreted as an indication that this was indeed the case, at this stage such a conclusion would be premature.)

One should have a close look at subpopulations of the labour force. For instance, using the subgroups considered in this and the previous section, one can decompose the gross wage rate densities for the whole population in two ways. Firstly, one can write each density as a convex combination ("superposition") of the densities for the subgroups "female workers" and "male workers". The densities for females and males can then be further decomposed by considering female (resp. male) manual and non-manual workers. Secondly, one may write the gross wage rate densities for the total population as mixtures of the densities for manuals and non-manuals, where the latter densities can, in turn, be expressed in terms of densities for female and male workers. In both decompositions we are faced with four subdistributions which have to be investigated very carefully if one wants to properly understand the changes in the aggregate distributions as shown in Figure 12. Possibly, one should also estimate separately densities for full-time and part-time workers.

Of course, most satisfactory would be a formal test of the hypothesis that the unknown densities which gave rise to the samples did not change over the 15 years under consideration. Devising tests of the hypothesis that two samples were drawn from the same distribution is an old, celebrated problem in statistics. So far it has been standard practice in the literature to examine the "two-sample" problem by either calculating a contingency table and performing a χ^2-test (see Kendall and Stuart, 1973, p. 576), or by using a test statistic based on the empirical cumulative distribution function (see Darling, 1957, for a useful survey). But it is also possible to construct a confidence band around a kernel estimator (see Härdle and Jerison, 1988, for details). If a density estimate for another year is contained in the confidence band, then one cannot reject (at the given level of significance) the hypothesis that the two corresponding samples were drawn from the same distribution; if the null hypothesis is violated, the test shows how strongly it is violated.

In the next chapter we need the gross wage rate densities in order to compute the elasticity of per capita labour supply (resp. per capita net earnings) with respect to the wage level. The relevant question within the present analysis is therefore whether the observed shifts in the normalised gross wage rate distributions will have a significant influence on the labour supply (resp. net earnings) elasticity. As we will see, this is not the case.

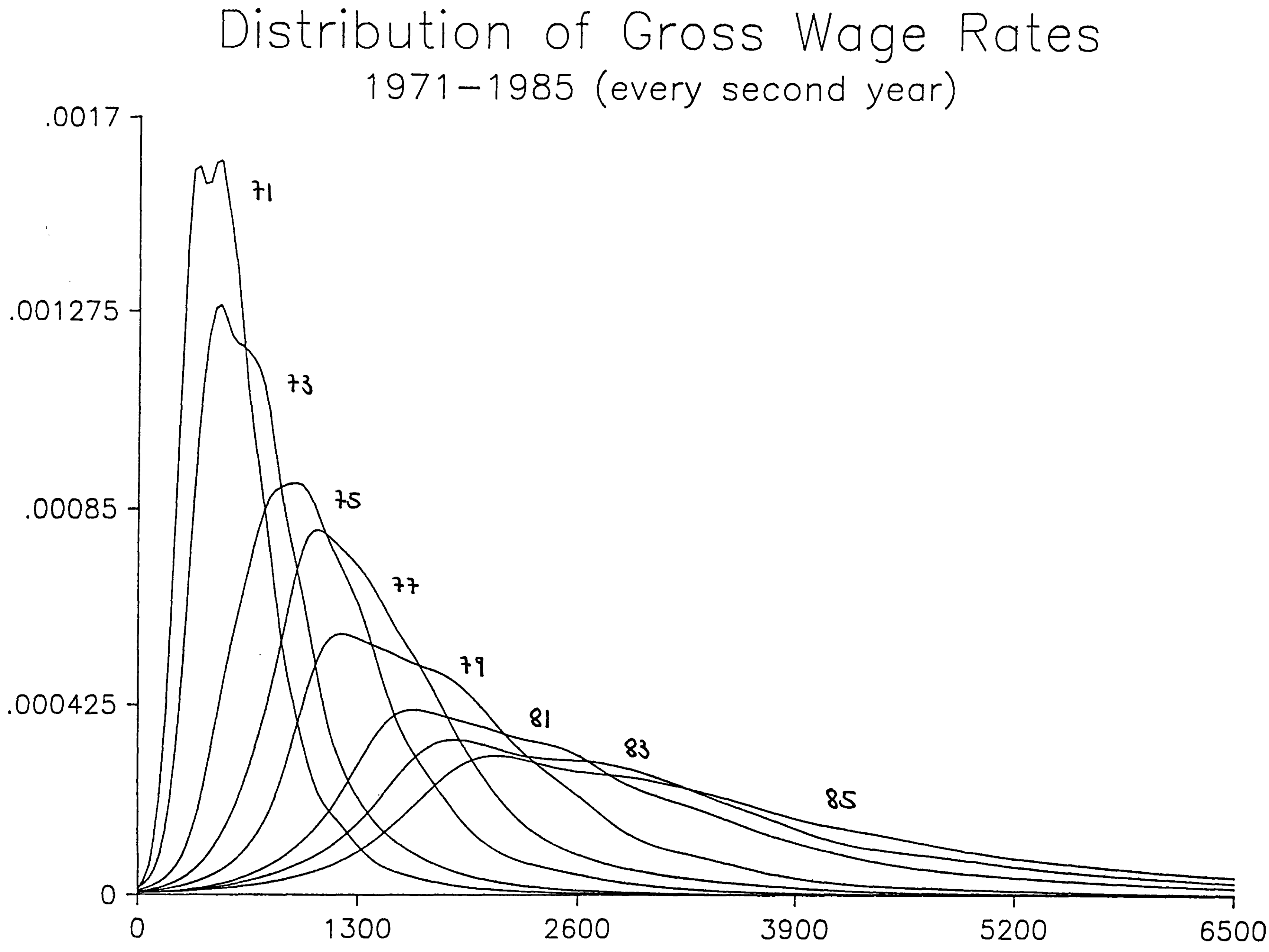
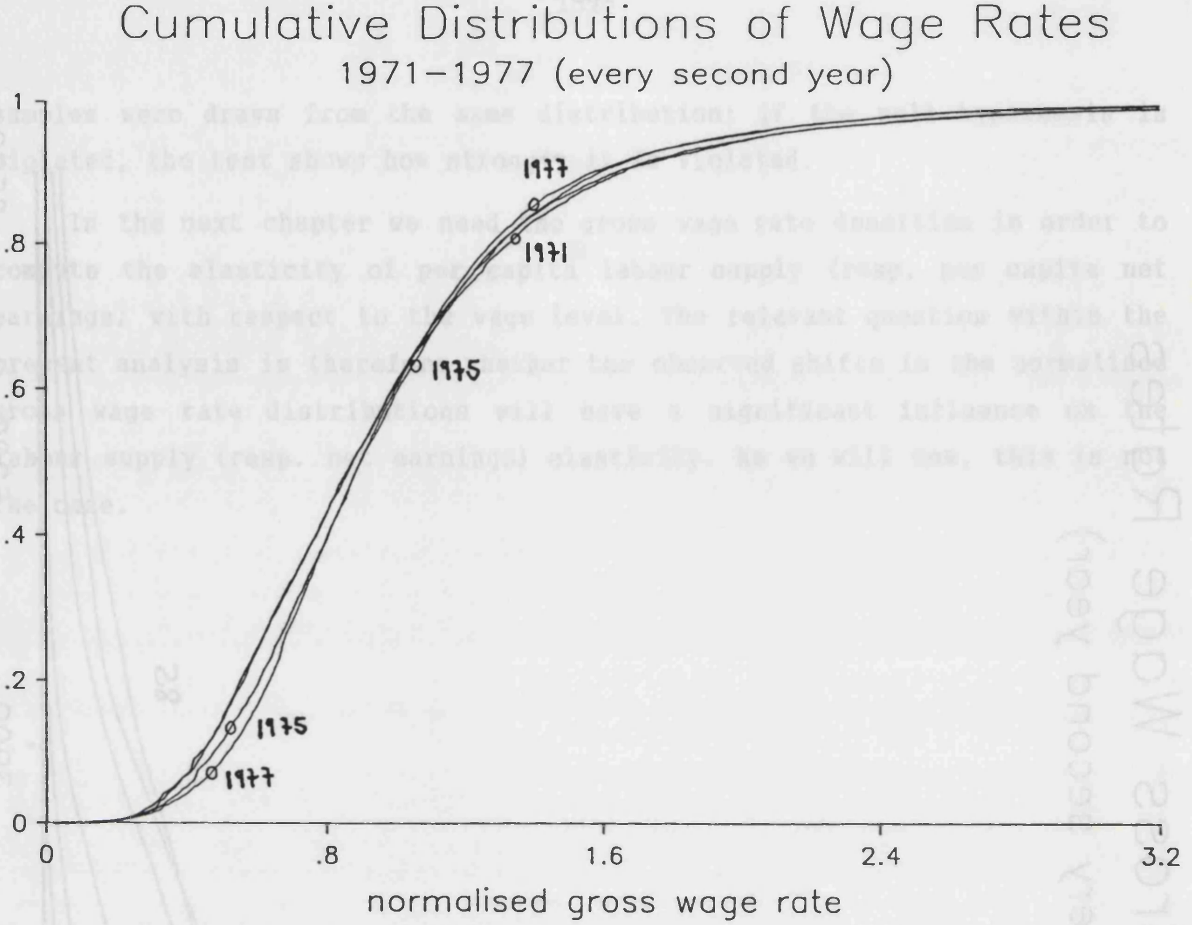


Figure 10 Adaptive kernel estimates


Cumulative Distributions of Wage Rates
1971–1977 (every second year)
1
.8
.6
.4
.2
0
0
.8
1.6
2.4
3.2
normalised gross wage rate
1977
1971
1975
1975
1977

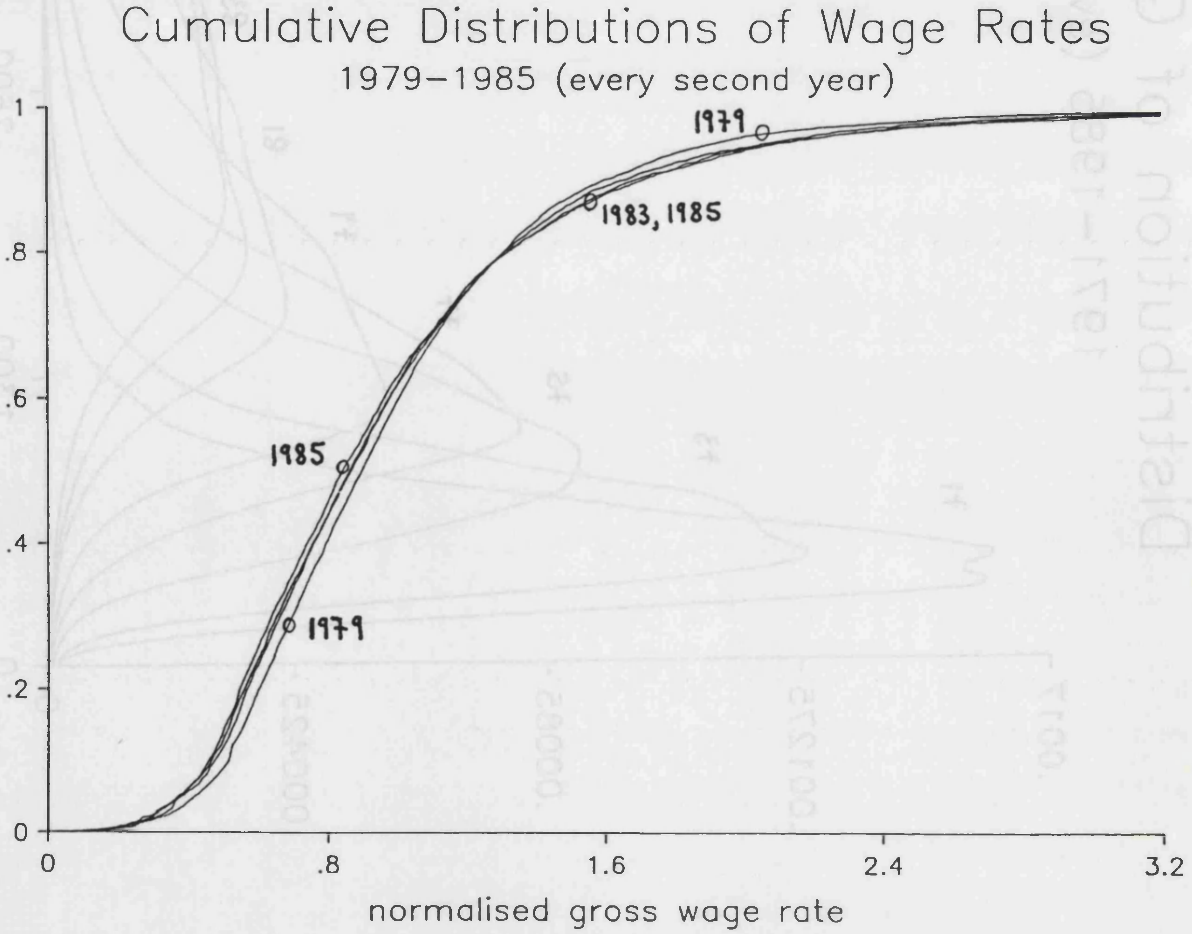
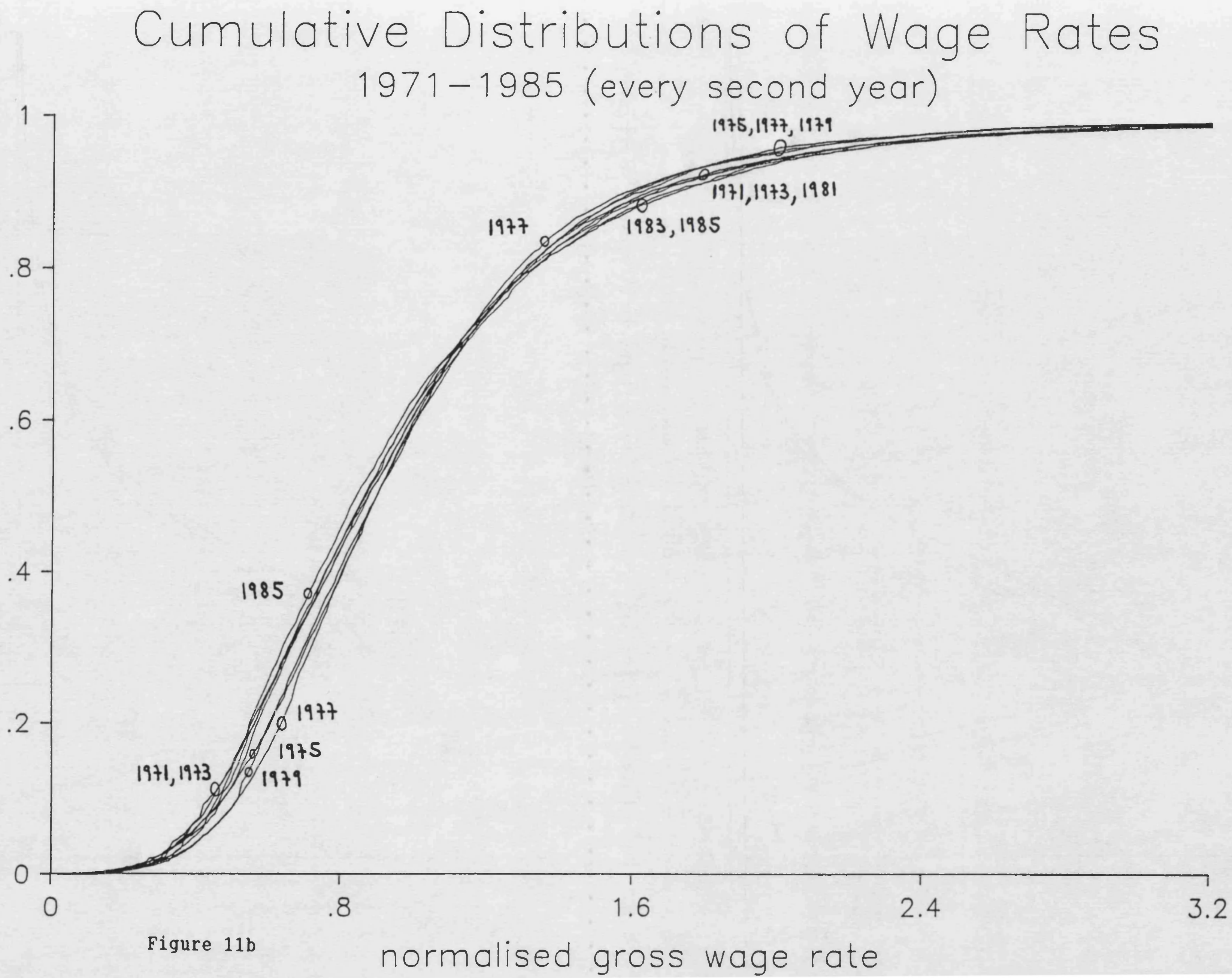
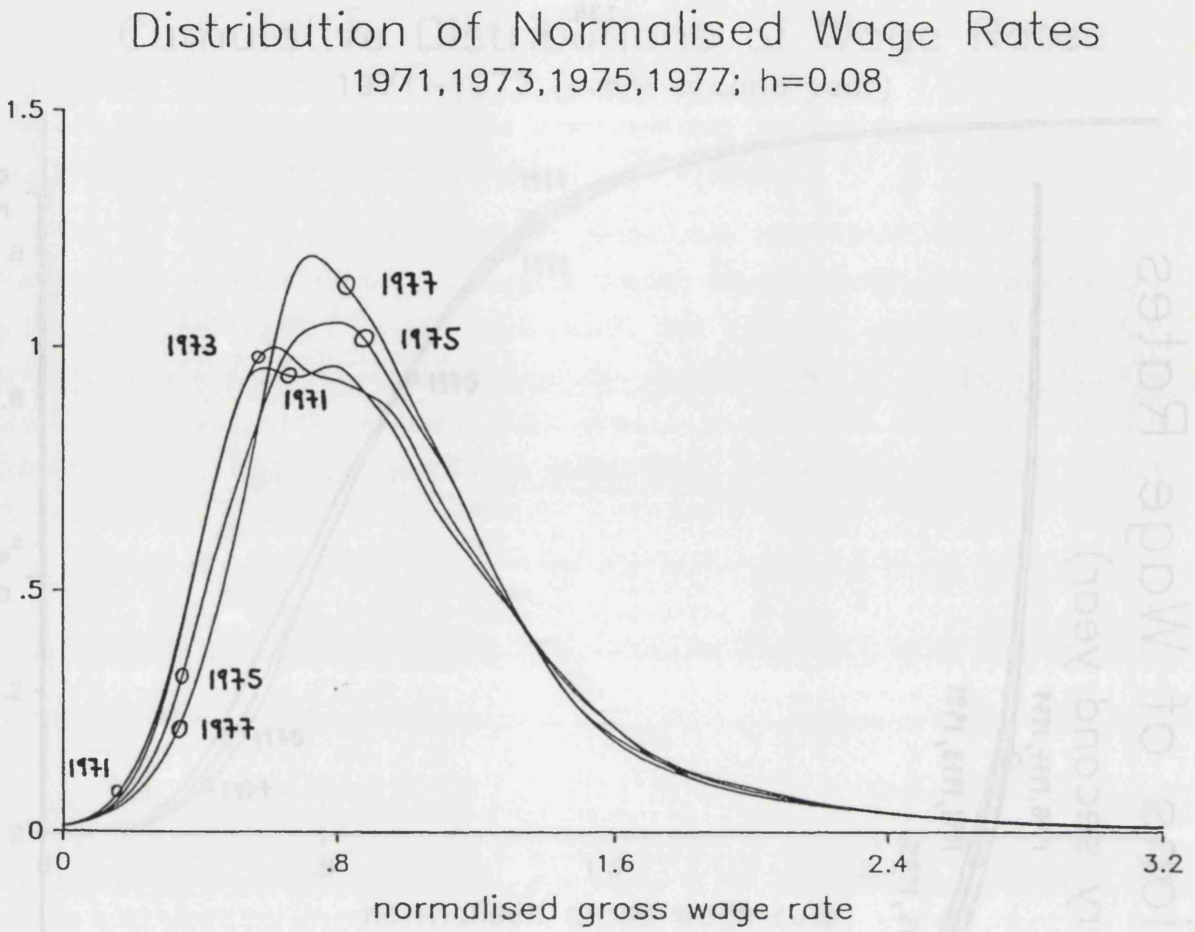

Cumulative Distributions of Wage Rates
1979–1985 (every second year)
1
.8
.6
.4
.2
0
0
.8
1.6
2.4
3.2
normalised gross wage rate
1979
1983, 1985
1985
1979

Figure 11a



Cumulative Distributions of Wage Rates
1971–1985 (every second year)
1975, 1977, 1979
1971, 1973, 1981
1977
1983, 1985
1985
1977
1975
1971, 1973
1979
0
.2
.4
.6
.8
1
0
.8
1.6
2.4
3.2
normalised gross wage rate

Figure 11b



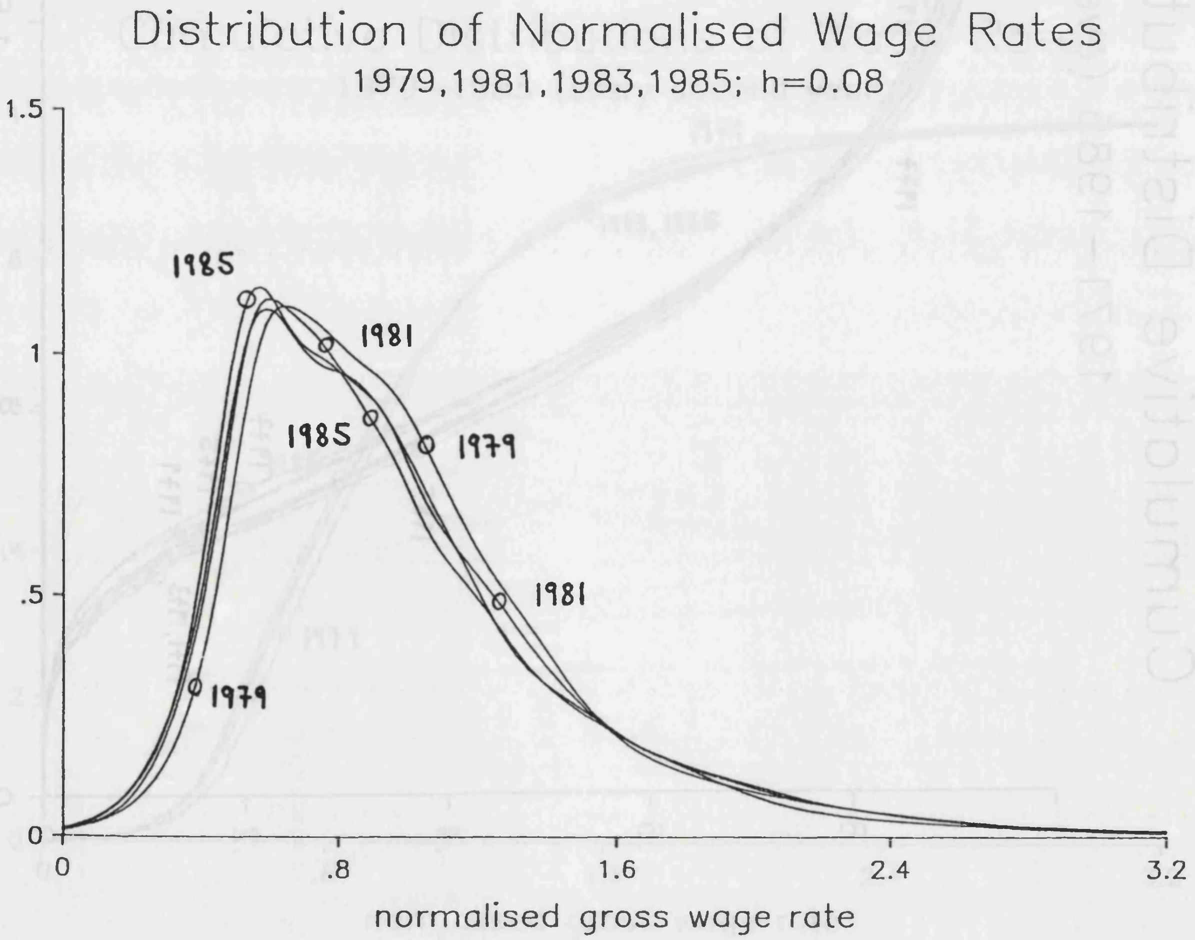


Figure 12a Adaptive kernel estimates

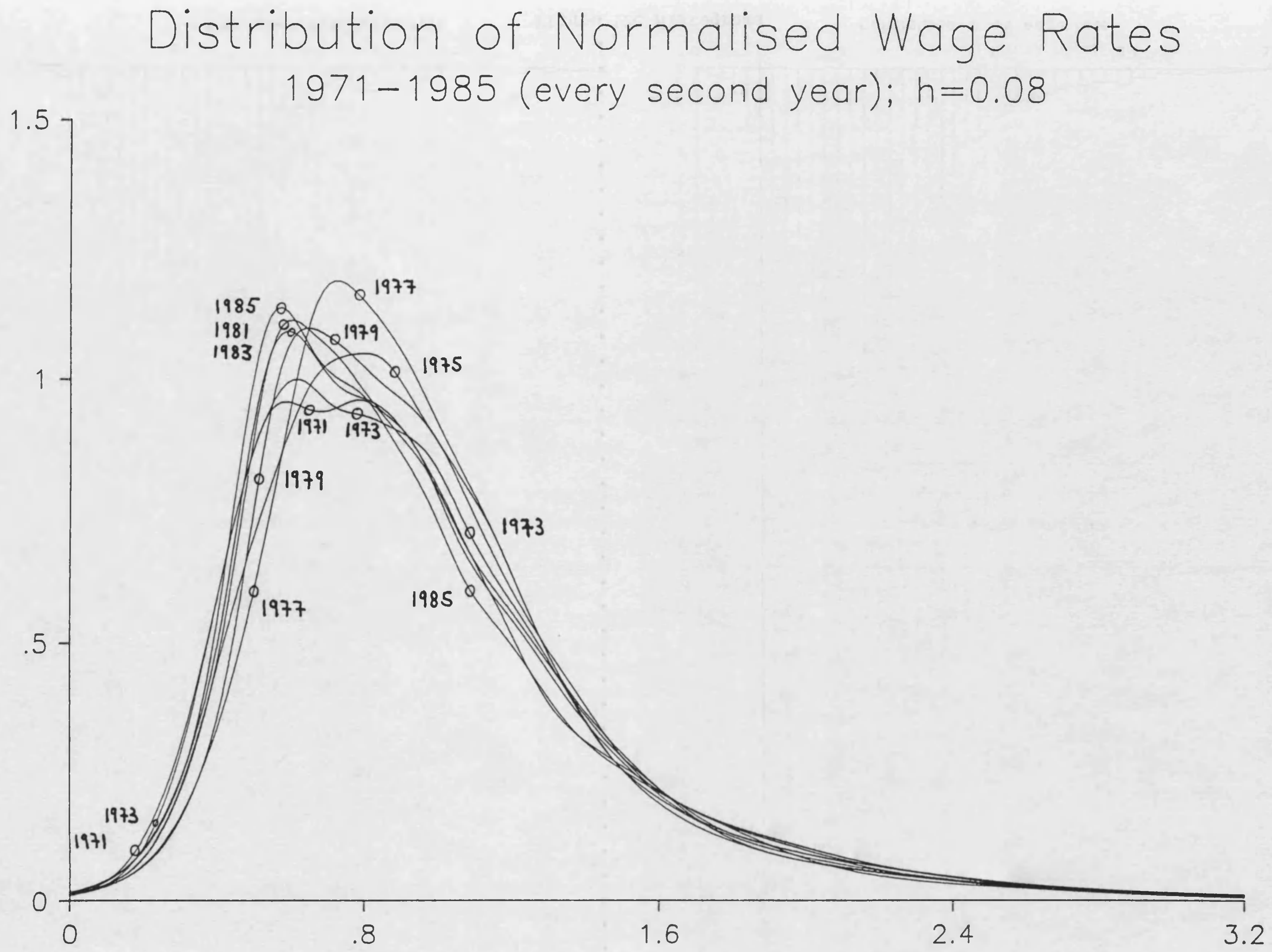


Figure 12b
Adaptive kernel estimates

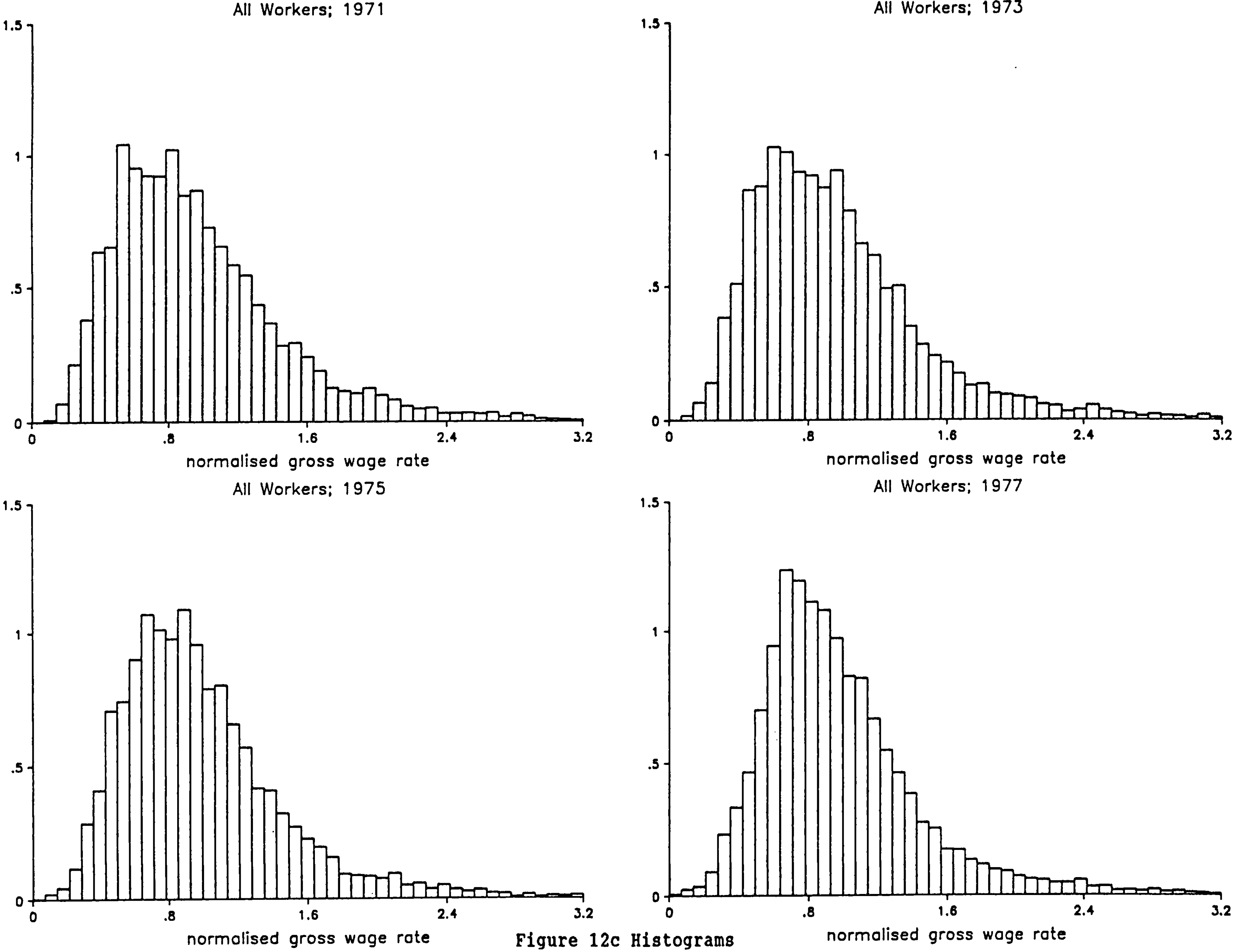


Figure 12c Histograms

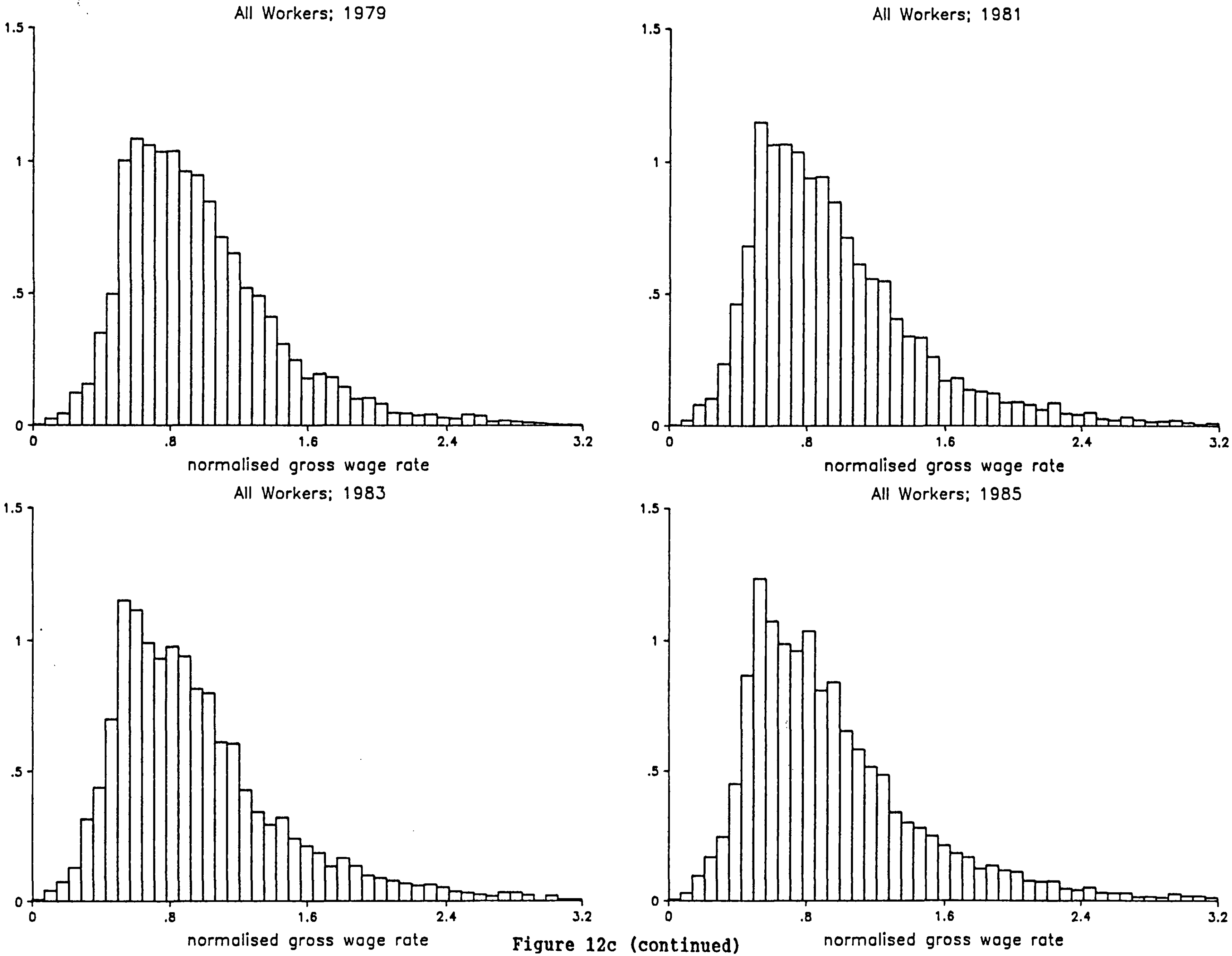
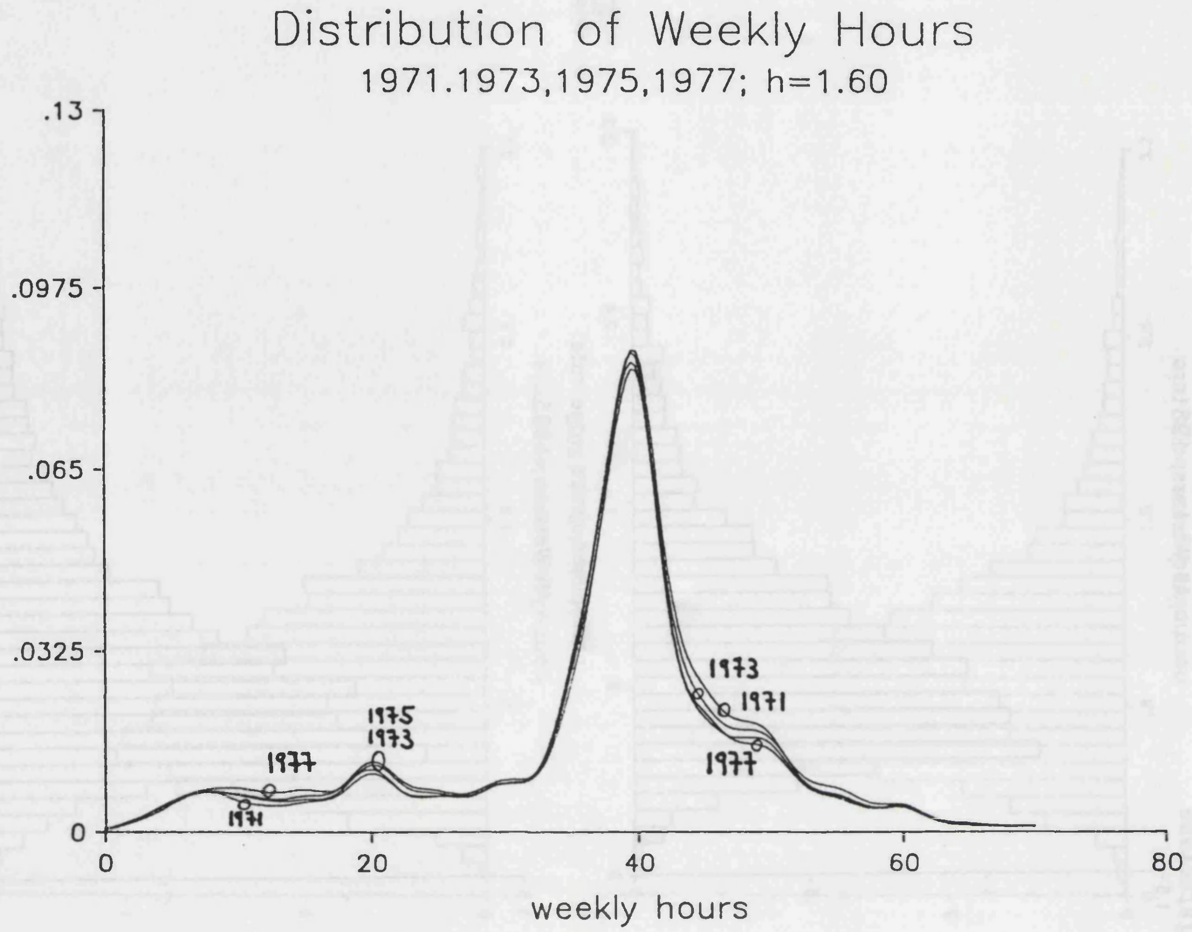


Figure 12c (continued)



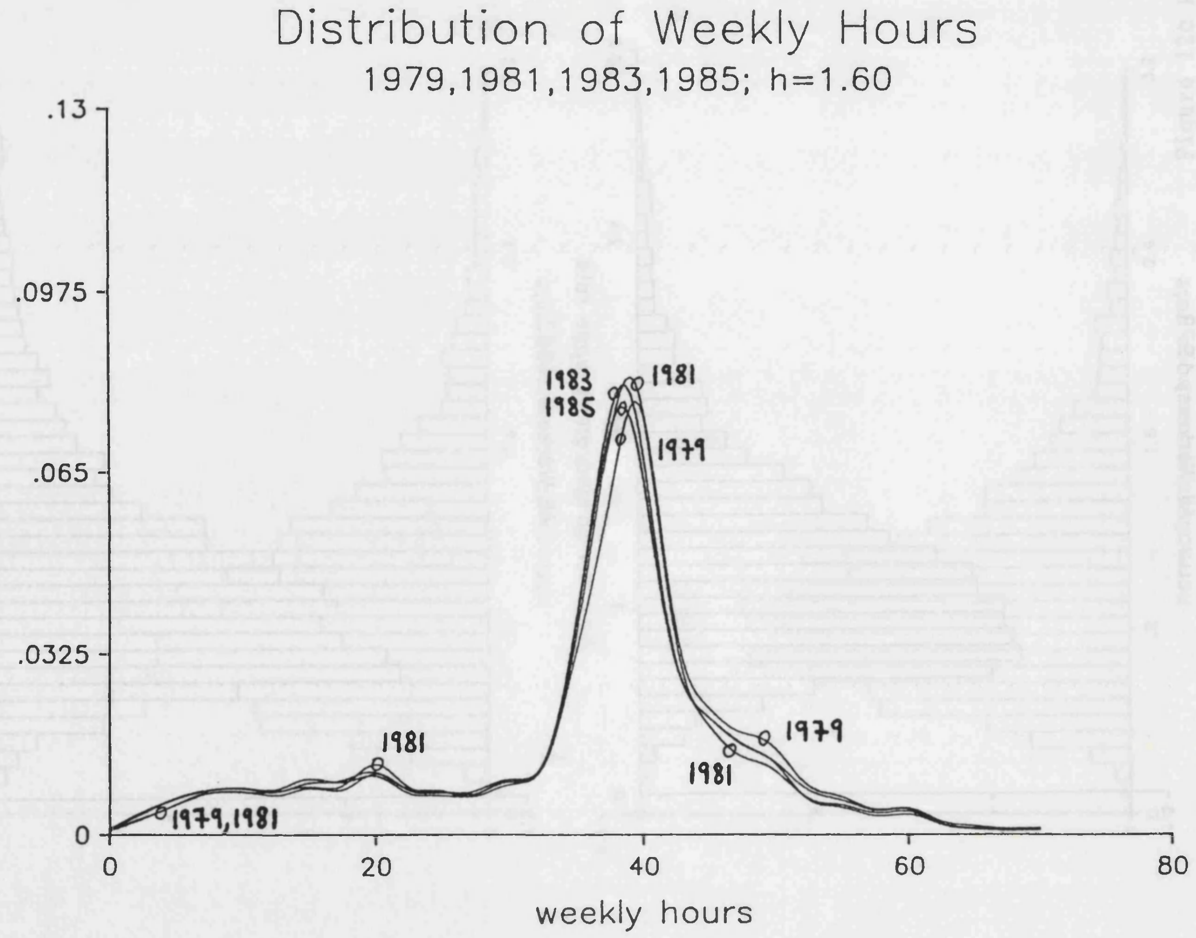
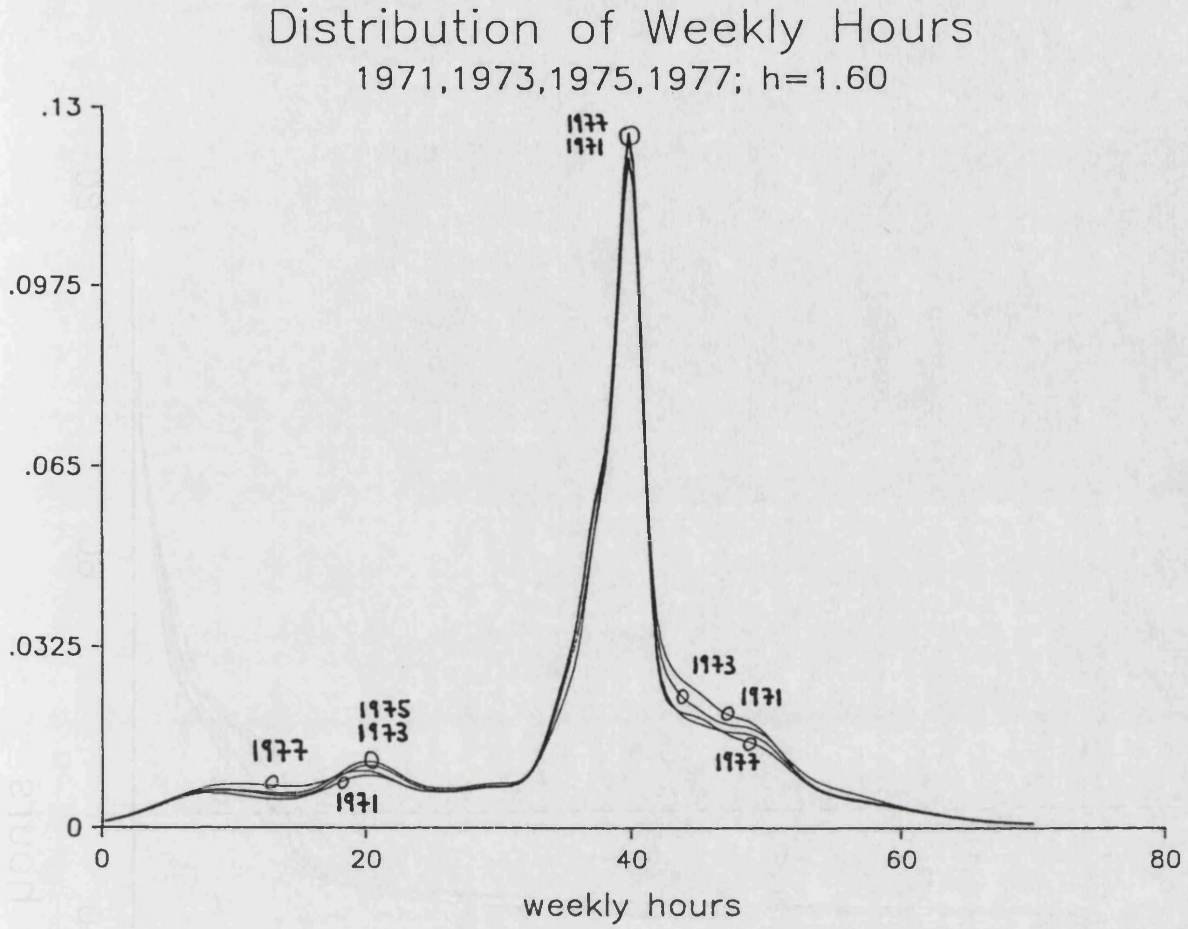


Figure 13a Ordinary kernel estimates



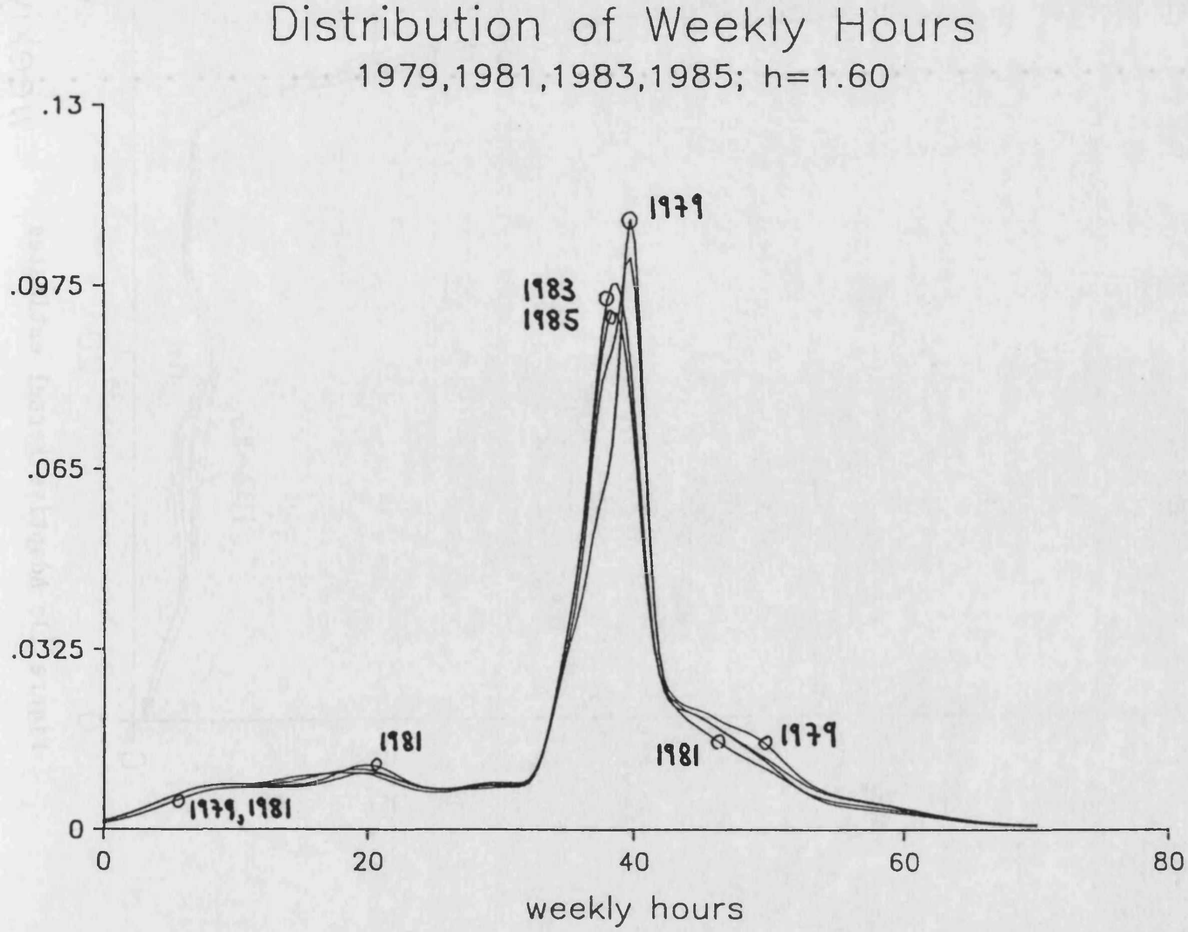


Figure 13b Adaptive kernel estimates

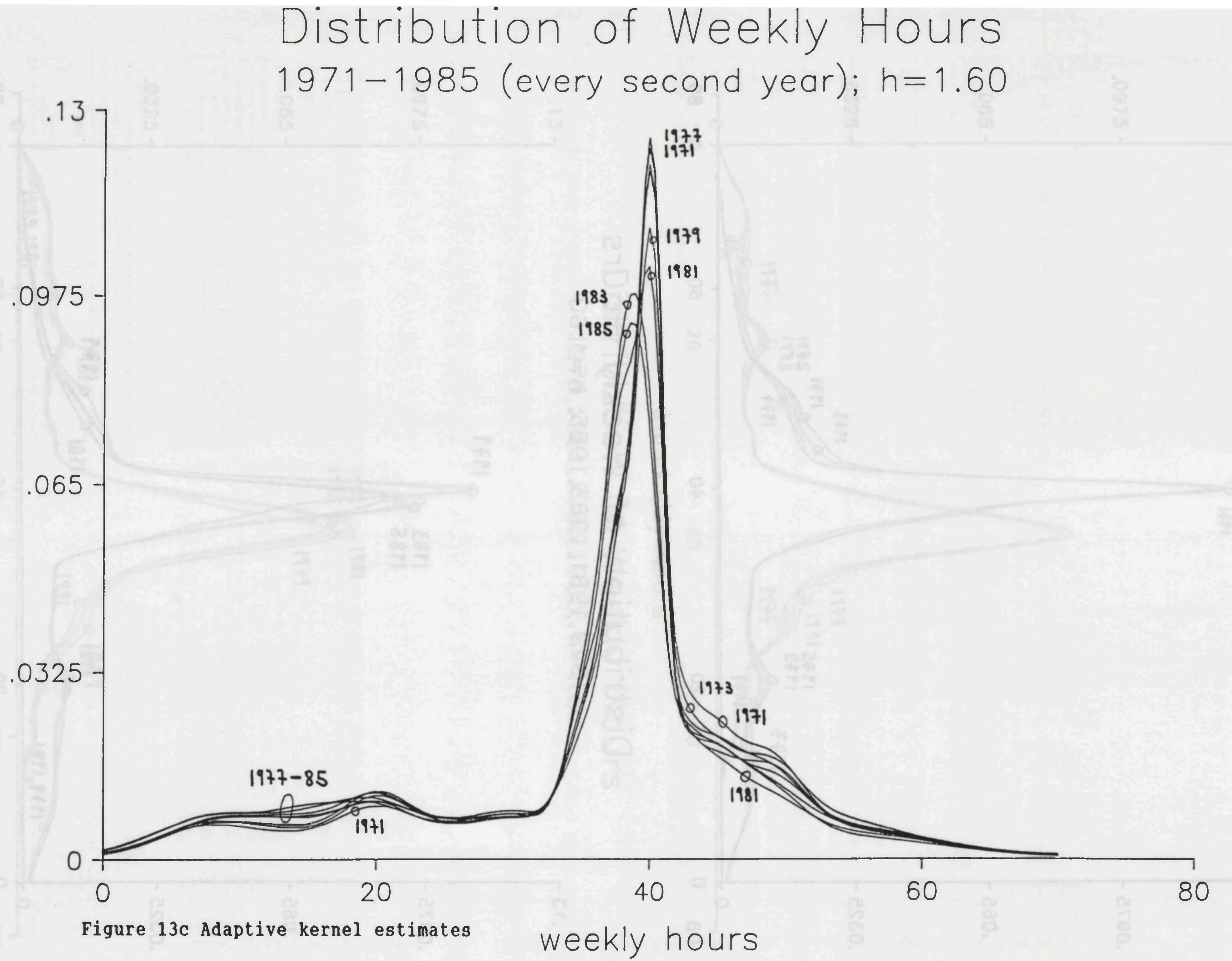


Figure 13c Adaptive kernel estimates

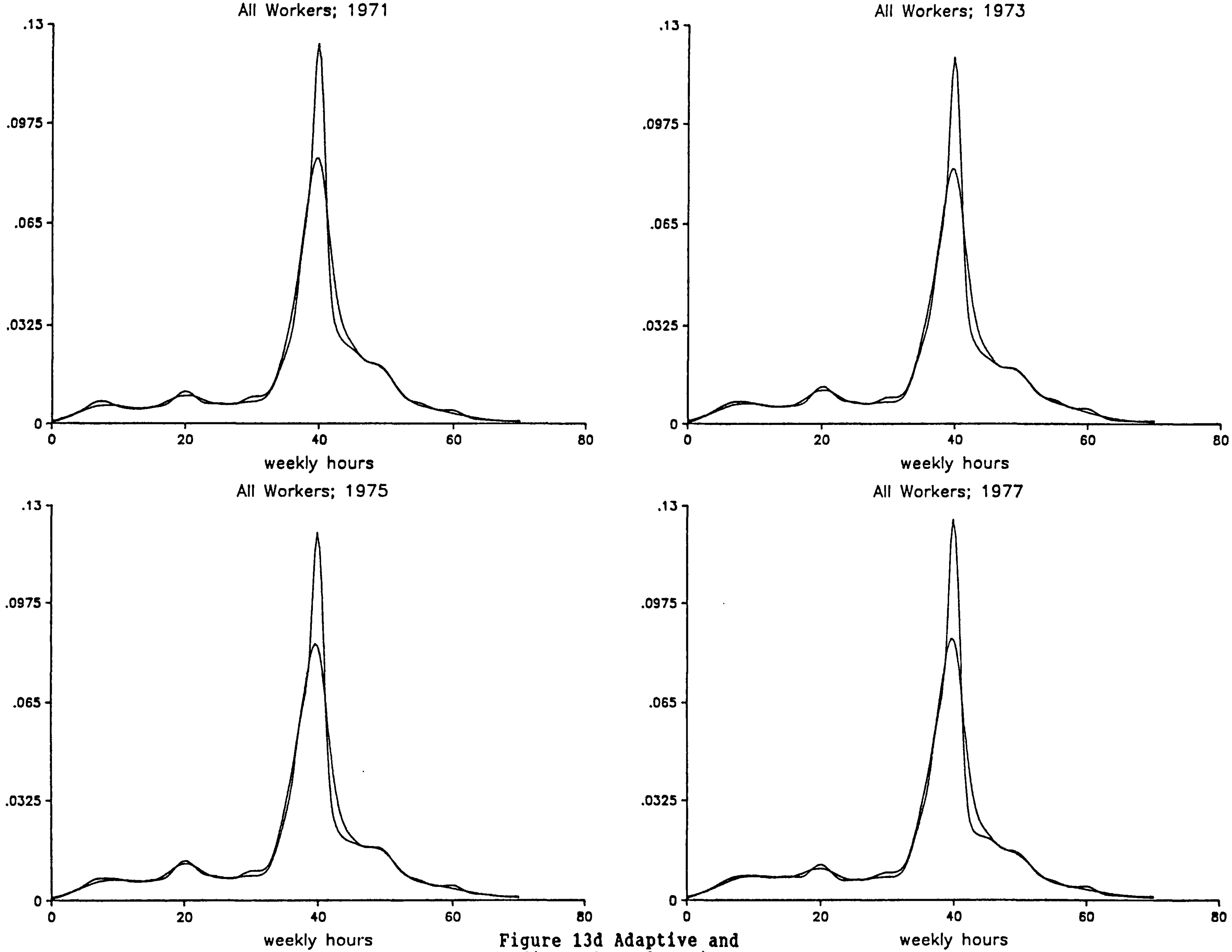


Figure 13d Adaptive and ordinary kernel estimates

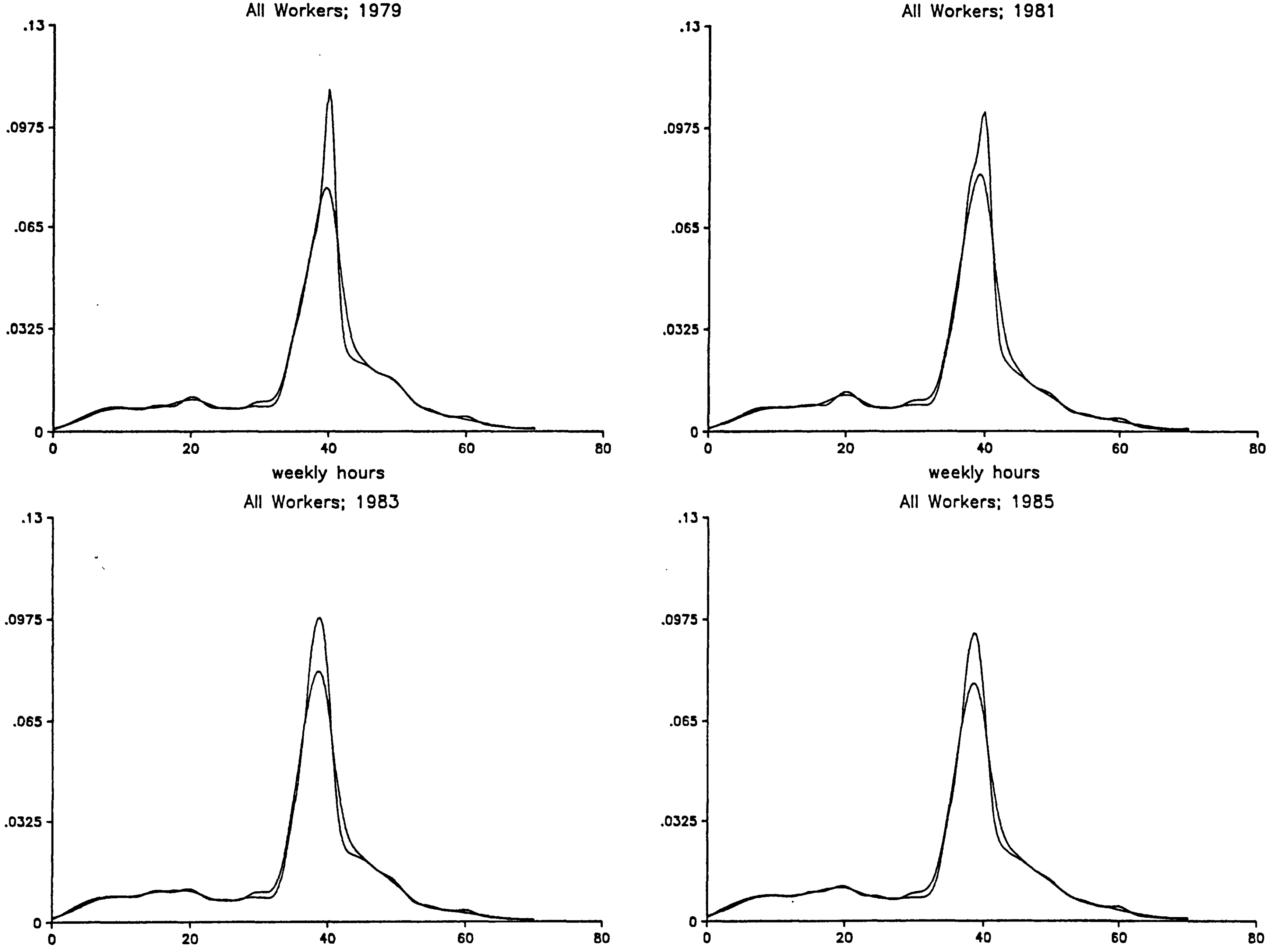


Figure 13d (continued)

6. Tests of Lognormality

We mentioned already in the Introduction that until now not very much is known about the determinants of the distributions considered in this chapter. (In fact, one reason why we applied nonparametric methods to the FES data was that a reliable knowledge of the form of these distributions may be helpful for developing a theory of the distribution of wages in a market economy.) On the other hand, economists typically make strong distributional assumptions in their empirical work.[12] Accordingly, there is an extensive econometric literature on curve estimation within a parametric framework (see the next chapter for references). Well-known functional specifications for univariate distributions are, e.g., the lognormal, beta and gamma distribution [see, e.g., Johnson and Kotz (1970a, 1970b) for a discussion of these and other distributions]. Especially the lognormal distribution has received much attention in the literature. Possibly many economists would agree that the size distribution of earnings is approximately lognormal, but with an upper tail which is better described by the density of the Pareto distribution.

Perhaps the best known work in this area is that of Lydall (1968). Lydall examined earnings data for a large number of countries and concluded that in general "the central part of the distribution from perhaps the tenth to the eightieth percentile from the top is close to lognormal" (p. 67). He argues, however, that there are more people with high earnings than the lognormal distribution would predict. In fact, the upper tail would follow more closely the Pareto distribution. In order "to account for the Pareto distribution of higher salaries" (pp. 127-128) he advances in chapter 4 of his book a "model of hierarchical earnings" based on the notion that "large organisations - which dominate the upper tail of the distribution - are organised on a hierarchical principle" (p. 9).[13]

Distributional assumptions of this type are in general not very well supported by the data. The null hypothesis, that the unknown density belongs to a known parametric class of distributions, is by standard tests very often rejected by the data.[14] This is, of course, disappointing. On the other hand, a functional form which has been rejected by a goodness-of-fit test may still provide a "reasonable" approximation to the data.

After having obtained an impression of the FES data, it is therefore interesting to compare the actual observed distributions with those generated by a standard distributional model. Since the lognormal distribution has received very much attention in the literature, we tested for all years from 1970 until 1985 and for several groups of workers the hypothesis that the data stem from a lognormal distribution. The variables chosen were: gross wage rate, gross earnings, net earnings and weekly hours of work. The test statistic employed was the Kolmogorov D-statistic which is based on a comparison of the empirical with the hypothetical distribution function.

We first considered the annual samples of "all workers" and the following eight subsamples: males, manual males, non-manual males, females, manual females, non-manual females, manuals and non-manuals. We then excluded from the data sets those workers who stated in the FES questionnaires that they worked usually not more than 30 hours per week and tested the null hypothesis for the nine samples of full-time workers.

The density of the two-parameter lognormal distribution is given by

$$\rho(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\cdot\frac{1}{x}\cdot\exp\{-\frac{(\log x-\mu)^2}{2\sigma^2}\}, \quad x > 0,$$

where μ and σ^2 are the parameters to be estimated. The maximum likelihood estimators of the two parameters are

$$\hat{\mu} = \frac{1}{n}\cdot\Sigma\log x_i$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\cdot\Sigma(\log x_i - \hat{\mu})^2,$$

where $x_1,\ldots,x_n$ are the observations (the standard reference for the lognormal distribution is Aitchison and Brown, 1957; but see also Johnson and Kotz, 1970a, Chapter 14). Figure 14 on the next page compares a lognormal maximum likelihood estimate of the distribution of gross wage rates for the year 1983 with an adaptive kernel estimate (computed with h=250). Obviously, the lognormal distribution does not fit the sample very well; and as we will see below, the tests rejected lognormality for the whole population in all years. The maximum likelihood estimates of μ and σ^2 are: $\hat{\mu}$=7.921 and $\hat{\sigma}^2$=0.285; hence the mean, the standard deviation and the skewness of the estimated lognormal distribution are, respectively, 3.177, 1.825 and 1.913, (in £; see also note 4). The values of the corresponding sample statistics are 3.184, 1.979 and 3.849, respectively.

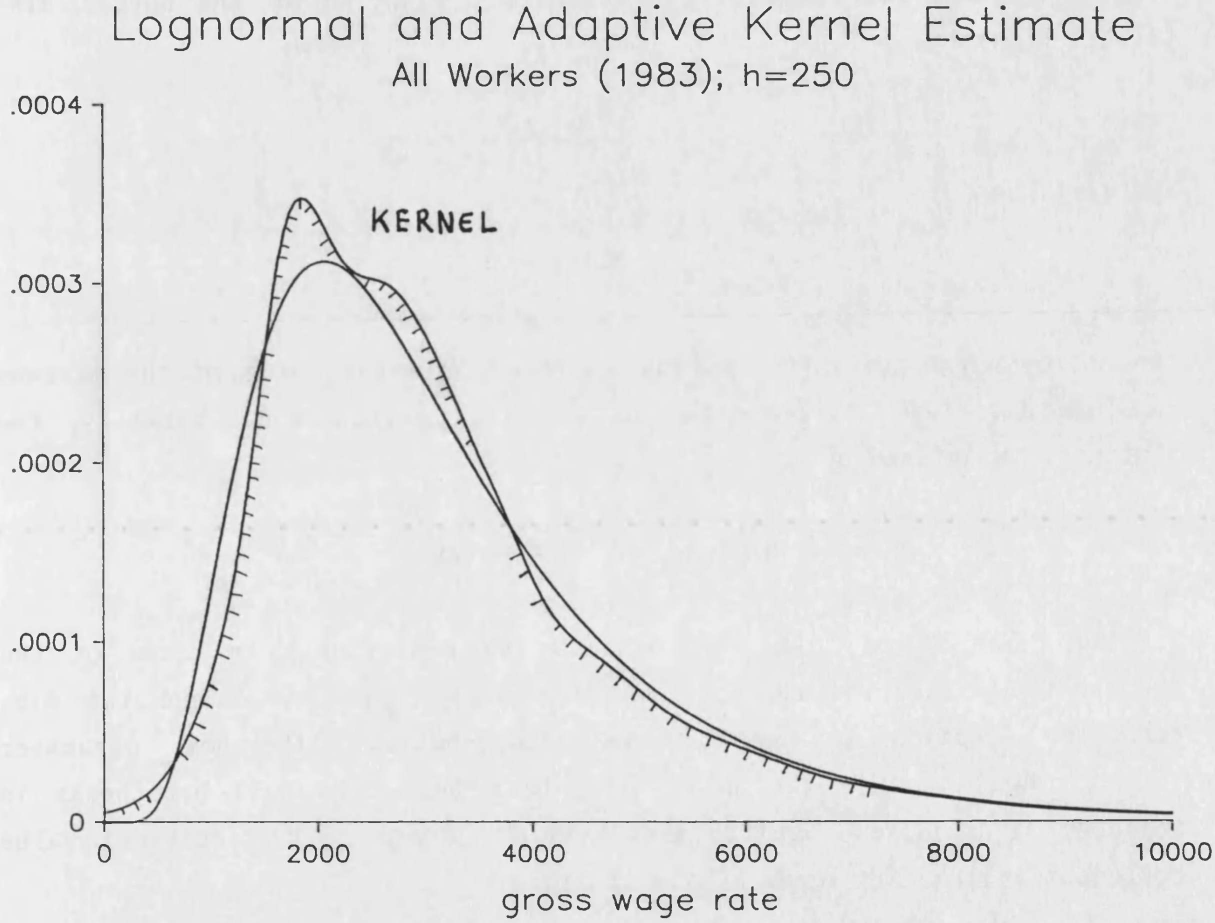


Figure 14

Let us now turn to the goodness-of-fit tests. The observations are treated as if they were realisations of independent and identically distributed random variables. Setting $y_i = \log x_i$ $(i=1,\ldots,n)$, we test the (composite) hypothesis that the observations $y_1,\ldots,y_n$ were drawn from a normal distribution. The mean and the variance of the normal distribution (i.e., the parameters μ and σ^2 of the lognormal distribution) are estimated by the sample mean and the sample variance. Let F_n and F_0 denote, respectively, the empirical cumulative distribution function of the sample $(y_1,\ldots,y_n)$ and the cumulative distribution function of the normal distribution with mean

$$\mu = \frac{1}{n}\cdot\Sigma y_i ,$$

and variance

$$\sigma^2 = \frac{1}{n-1}\cdot\Sigma(y_i - \mu)^2 .$$

The Kolmogorov D-statistic assigns to the sample the value of the maximum absolute difference between the two distribution functions. Formally, the statistic is defined by

$$D = \sup |F_n(y) - F_0(y)| .$$

Setting $y=\log x$, one can also express the D-statistic in terms of the empirical distribution function of the sample $(x_1,\ldots,x_n)$ and the distribution function of the lognormal distribution with above parameter values. The D-test is the usual upper-tail test. The null hypothesis is rejected at a chosen significance level if D exceeds the critical value corresponding to this level of significance.[15)]

The greater the sample size, the less likely is a given deviation between F_n and F_0 (if the null hypothesis is true). To carry out the tests, the SAS procedure "Univariate" was used; the programme obtains the probability that D assumes a larger value than that observed - in the following denoted Prob(D>d) - by calculating the modified statistic

$$D^* = (\sqrt{n} - 0.01 + 0.85/\sqrt{n})\cdot D$$

and interpolating linearly within the range of simulated percentage points for D* given by Stephens (1974):

15%	10%	5%	2.5%	1%
0.775	0.819	0.895	0.955	1.035

The null hypothesis was rejected by the D-test in most cases. Let us begin with those samples which contain both full-time and part-time workers. Obviously, the data do not support the hypothesis of a lognormal distribution of weekly hours of work (see Figures 3, 8 and 13). Regardless of which sample was considered, the null hypothesis was always decisively rejected. Looking at the earnings data, there are only three samples which possibly could have been generated by a lognormal distribution, namely the samples for males, manual males and non-manual males (see Figures 4 and 9). But again, the null hypothesis was in all years rejected by the data at the one per cent level.

Table 9 below contains the values of the D-statistic for the gross wage rate data. Notice that the D-statistic assumes its smallest values on the samples of non-manual workers; the statistic's values for manual workers are in all years much larger. The empirical distribution functions deviate most strongly from the hypothetical distribution functions in the case of manual female workers. In the years 1970-75 the deviations are larger for male workers than for female workers, while from 1976 onwards they are larger for females. Reading down the columns for non-manual female and non-manual male workers, there are again only 6 years where the deviation from the null hypothesis is smaller for females. The null hypothesis was accepted at the 5 per cent level for the subgroup "non-manual workers" in the years 1973, 1974 and 1975; the corresponding probabilities of observing a larger value of the D-statistic are 0.077, 0.115 and 0.058, respectively. For all other years and subgroups the alternative (that the data do not come from a lognormal distribution) was significant at the one per cent level. The pragmatist may argue, however, that many of the observed deviations are "acceptable".

Table 9 Tests of Lognormality
Values of the Kolmogorov D-statistic; variable:
gross wage rate; "all workers" and subsamples

YEAR	ALL WORKERS	MEN	WOMEN	MANUALS	NON-MAN.	MANUAL MEN	MANUAL WOMEN	NON-MAN. MEN	NON-MAN. WOMEN
1970	0.019	0.058	0.057	0.052	0.025	0.078	0.077	0.038	0.054
1971	0.018	0.054	0.050	0.046	0.017	0.076	0.066	0.041	0.037
1972	0.021	0.055	0.051	0.050	0.017	0.081	0.063	0.034	0.034
1973	0.024	0.058	0.058	0.047	0.014 1)	0.075	0.077	0.041	0.038
1974	0.022	0.051	0.048	0.049	0.014 2)	0.074	0.084	0.036	0.033
1975	0.024	0.054	0.043	0.055	0.014 3)	0.081	0.079	0.027	0.032
1976	0.037	0.060	0.066	0.062	0.023	0.086	0.089	0.032	0.060
1977	0.035	0.049	0.063	0.050	0.020	0.068	0.107	0.028	0.053
1978	0.034	0.049	0.064	0.046	0.019	0.063	0.103	0.037	0.053
1979	0.032	0.043	0.076	0.035	0.022	0.064	0.119	0.036	0.051
1980	0.024	0.050	0.056	0.035	0.019	0.071	0.092	0.040	0.043
1981	0.027	0.041	0.060	0.034	0.017	0.064	0.096	0.038	0.040
1982	0.026	0.047	0.061	0.030	0.020	0.072	0.096	0.054	0.041
1983	0.031	0.048	0.067	0.036	0.025	0.075	0.111	0.048	0.049
1984	0.029	0.048	0.060	0.033	0.022	0.081	0.107	0.040	0.041
1985	0.031	0.038	0.066	0.038	0.020	0.063	0.123	0.043	0.041

1) P(D>d)=0.077; 2) P(D>d)=0.115; 3) P(D>d)=0.058

Table 10 Tests of Lognormality

Values of the Kolmogorov D-statistic; variable: gross wage rate; full-time workers; P(D>d) in brackets

YEAR	ALL F.-T. WORKERS	MEN	WOMEN	MANUALS	NON-MAN.	MANUAL MEN	MANUAL WOMEN	NON-MAN. MEN	NON-MAN. WOMEN
1970	0.032	0.050	0.034	0.063	0.016 (>0.15)	0.064	0.069	0.034	0.020 (>0.15)
1971	0.033	0.049	0.039	0.060	0.011 (>0.15)	0.062	0.079	0.037	0.029 (=0.019)
1972	0.033	0.046	0.032	0.060	0.017 (=0.043)	0.064	0.067	0.028	0.035
1973	0.042	0.056	0.033	0.070	0.013 (>0.15)	0.068	0.087	0.038	0.020 (>0.15)
1974	0.030	0.044	0.043	0.055	0.014 (>0.15)	0.058	0.076	0.029	0.037
1975	0.034	0.044	0.039	0.066	0.015 (=0.126)	0.067	0.084	0.022 (=0.043)	0.020 (>0.15)
1976	0.045	0.054	0.051	0.068	0.015 (>0.15)	0.073	0.093	0.026 (=0.014)	0.044
1977	0.041	0.043	0.069	0.057	0.022	0.056	0.122	0.024 (=0.032)	0.051
1978	0.036	0.041	0.066	0.059	0.016 (=0.095)	0.050	0.116	0.034	0.044
1979	0.027	0.037	0.050	0.049	0.015 (=0.14)	0.051	0.099	0.035	0.032
1980	0.028	0.042	0.045	0.052	0.017 (=0.038)	0.056	0.084	0.031	0.036
1981	0.026	0.034	0.055	0.050	0.016 (=0.054)	0.049	0.107	0.030	0.038
1982	0.031	0.041	0.046	0.053	0.015 (=0.13)	0.056	0.092	0.048	0.033
1983	0.033	0.040	0.062	0.057	0.021	0.062	0.102	0.037	0.051
1984	0.027	0.040	0.043	0.054	0.017 (=0.051)	0.064	0.098	0.035	0.036
1985	0.028	0.034	0.056	0.046	0.018 (=0.038)	0.054	0.119	0.037	0.042

It is interesting to observe that the null hypothesis was more supported by the gross wage rate data for full-time workers. Table 10 above contains the values of the D-statistic for the nine groups of full-time workers. In those cases where the alternative was not significant at the 1 per cent level, the probability of observing a larger test statistic is given in brackets (but recall that the percentage points for the D-statistic are simulated and not exact values). Notice first that the pattern of deviations (of the empirical from the hypothetical distribution functions) is the same as that in Table 9.

Reading down the column for full-time non-manual workers, we see that the null hypothesis was accepted for this subgroup at the 1 (resp. 5) per cent level in 14 (resp. 11) of the 16 years. Moreover, the probability of observing a larger test statistic is in many years much higher than 5 per cent: in 8 (resp. 5) years it is greater than 12 (resp. 15) per cent. Thus, the FES gross wage rate data for the subgroup "full-time non-manual workers" support the null hypothesis very well. As we will see below, we can accept lognormality for all years from 1970 until 1985 at a significance level of 0.01.

For the remaining groups of workers the null hypothesis performed less well. Looking down the last two columns of the table, we see that for full-time non-manual male workers the null hypothesis was accepted at the 1 per cent level in the years 1975, 1976 and 1977; for full-time non-manual female workers it was accepted at the 1 per cent level in 1971 and at the 15 per cent level in 1970, 1973 and 1975. For all other years and subgroups the values of the D-statistic were significant at the 1 per cent level.

After having excluded part-time workers from the samples, the tests still showed an extremely poor performance of the lognormal function as a description of the labour supply distributions. The null hypothesis was in all cases strongly rejected by the data. Clearly, this was to be expected. To give an example, consider full-time workers in manual occupations. As we see in Table 10, the D-statistic for the gross wage rate data assumes in 1971 the value 0.060. In the case of the labour supply data, we have D= 0.226; if we include also part-time workers, we obtain D=0.319.

Turning to gross earnings, we found that the null hypothesis was again surprisingly well supported by the data for the subgroup "full-time non-manual workers". As the figures in Table 11 on the next page show, the hypothesis was accepted in 11 (resp. 6) of the 16 years at the 1 (resp. 5) per cent level.

Table 11 Further Tests of Lognormality

Cases of non-rejection; full-time non-manual workers; variable: gross weekly earnings

YEAR	D-STATISTIC	PROB(D>d)
1970	0.024	0.016
1974	0.013	>0.150
1975	0.018	0.023
1976	0.017	0.054
1978	0.015	>0.150
1979	0.019	0.019
1980	0.014	>0.150
1981	0.016	0.079
1983	0.019	0.022
1984[1)]	0.015	0.142
1985	0.018	0.030

1) 1984 also: net weekly earnings, D=0.020 and Prob(D>d)=0.013.

This good showing of the null hypothesis is surprising since it conflicts with the widely accepted "wisdom" on the size distribution of earnings mentioned at the outset of the section, namely that the lognormal function can only properly describe the main part of the earnings distribution, while its upper tail would follow more closely the Pareto law. We remark that Lydall's model of "hierarchical earnings", advanced to account for the Pareto upper tail, is an attempt to explain earnings differentials among non-manual workers. Consequently, if there was clear empirical evidence for the view that the lognormal function does not provide a good description of the upper range of the earnings distribution and, in particular, for the explanation of the upper tail given by Lydall, the D-tests should have strongly rejected the null hypothesis for the subgroup "full-time non-manual workers". It is interesting to observe that Harrison (1981), using another test statistic and the British New Earnings Survey data for 1972, also did not reject lognormality for this group of workers (see also note 14).

In the case of the net earnings data for full-time non-manual workers, the null hypothesis was rejected at the 1 per cent significance level in all years except 1984. Looking at the earnings data for the remaining groups of full-time workers, we do not find empirical evidence for lognormality. The null hypothesis was rejected at the 1 per cent level in all but the following cases:

Full-time manual females

1970, gross earnings: D=0.044, Prob(D>d)=0.026

Full-time non-manual females

1972, net earnings: D=0.029, Prob(D>d)=0.018
1973, net earnings: D=0.030, Prob(D>d)=0.014;
and gross earnings: D=0.018, Prob(D>d)>0.15

Full-time non-manual males

1975, net earnings: D=0.017, Prob(D>d)>0.15
1977, net earnings: D=0.020, Prob(D>d)=0.137;
and gross earnings: D=0.023, Prob(D>d)=0.037

So far we have been concerned with testing lognormality for single years. The question arises whether we can infer from the test results that from 1970 until 1985 all gross wage rate distributions for the subgroup "full-time non-manual workers" were lognormal. It is not difficult to answer the question. Suppose that this was indeed the case. Furthermore, suppose that we test the hypothesis in a single year at the significance level α. Let R denote the number of rejections in the 16 years under consideration. Recall that a deterministic α-test can be represented as a (0-1) random variable which assumes the value 1 if and only if the null hypothesis is rejected by the data; the probability that 1 occurs is equal to α. Thus, the random variable R has a binomial distribution with parameters $\alpha \in [0,1]$ and n=16. Accordingly, the probability of observing at least k rejections is given by

$$\text{Prob}(R \geq k) = \Sigma_{i \geq k} \binom{16}{i} \cdot \alpha^i \cdot (1-\alpha)^{16-i} \qquad (k=0,1,\ldots,16).$$

We will reject the hypothesis of lognormality in the years 1970-85 if Prob(R≥k) is smaller than a chosen level of significance. In the case of α=0.01 (resp. α=0.05) the null hypothesis was rejected in 2 (resp. 5) years. The probability of observing at least 2 (resp. 5) rejections is 0.011 (resp. 0.001) if α equals 0.01 (resp. 0.05). Thus, in the case of α=0.01, we can just accept at the 1 per cent level the hypothesis that the

16 gross wage rate samples for the subgroup "full-time non-manual workers" were drawn from lognormal distributions. Of course, the data do not support lognormality of the gross earnings distributions for this group in the period 1970-85. Recall that the null hypothesis was rejected at the 1 (resp. 5) per cent level in 5 (resp. 10) years. In both cases we have Prob(S≥k)=0, approximately.

The outcomes of the tests show how careful one has to be in drawing conclusions from samples which relate to a single year. Since a phenomenon observed in just one year is in general of little interest, one has to consider time-series of cross-sectional data in order to find out whether there is really empirical evidence for a distributional assumption. The test results for non-manual workers are very interesting. However, it would be premature to conclude that we have empirical evidence for a lognormal distribution of gross wage rates in the case of full-time workers in non-manual occupations. Firstly, the hypothesis was just accepted at the 1 per cent level. Secondly, we can only speak of empirical evidence if a phenomenon has been observed in several data sets and over a long time period.

In addition, the reader may have objections against the goodness-of-fit test employed here for testing lognormality. The maximum absolute difference between two functions is a very simple measure of deviation. It would be interesting to use a test statistic which measures in some sense the average deviation of the empirical from the theoretical distribution function.[16] In particular, it would be interesting to test distributional assumptions by constructing confidence bands around a kernel estimator. [This was done by Härdle and Jerison (1988) in the related field of regression estimation.]

The poor performance of the lognormal distribution in most cases is, of course, not astonishing. The hypothesis tested was not a sophisticated one, and future research in this area should pay attention to formulating and exploring distributional assumptions which go beyond simple hypotheses like that of lognormality.

7. Notes

1) The reader interested in theories of the personal distribution of incomes and earnings may find the surveys by Sahota (1978) and Atkinson (1979) useful (the latter is unfortunately only published in German); the articles collected in Atkinson (1976) pay attention to theory and evidence. Among the many papers written on the subject, an outstanding conceptual contribution is that by Friedman (1953) from which the quotation in the Introduction was taken. Thurow's (1976) remarks on the marginal productivity theory of distribution are worth reading. To account for the actual observed distribution of earned income, he advances a job competition mechanism based on the notion that workers compete for jobs rather than for wages (which are rigid and fixed by employers). The classic reference on the distribution of earnings is Lydall (1968); a recent contribution based on the human capital theory is the monograph by Weizsäcker (1987). A detailed discussion of the broader subject "inequality" is given by Atkinson (1984). It is interesting to observe that Atkinson concludes his book by writing: "...we have time and time again come to phenomenona for which no adequate explanation exists...One of the aims of the book has been to demonstrate that far too little is known about this central subject. This is an indictment of economics, but it is also a challenge" (pp. 284-285).

2) A natural extension of the next chapter's topic would be to estimate a complete system of commodity demands and labour supply. We remark that the expenditure data of the FES - especially expenditure on durables - are presumably less reliable than its income data: Since the expenditure data cover only a 14-day period, the FES records households who make purchases at intervals of more than 14 days as consuming either zero or more than their average amounts. The FES lacks also some of the variables which would be very useful for stratifying the annual samples of "all workers", such as the level of education and past work experience.

3) It is, of course, unsatisfactory to determine the smoothing parameter more or less by trial and error as we did in this work. We could have estimated h by least-squares or likelihood cross-validation. However, as Silverman (1986, p. 44) notes, "there is as yet no universally accepted approach to this problem".

4) It may be helpful to compare these figures with the skewness of some standard distributions. The following four distributions are skewed to the right and concentrated on R_+ (see, e.g., Johnson and Kotz, 1970a and 1970b, for precise definitions and proofs).

Lognormal distribution: The random variable $X>0$ is said to have the lognormal distribution with parameters μ and σ if log X is normally distributed with mean μ and standard deviation σ. We have: Mean = $\exp(\mu+\frac{1}{2}\sigma^2)$, Variance = $(\text{Mean})^2\cdot(\exp(\sigma^2)-1)$ and Skewness = $(\exp(\sigma^2)-1)^{1/2}\cdot(\exp(\sigma^2)+2)$. The skewness increases very rapidly with σ:

σ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Skewness	0.30	0.61	0.95	1.32	1.75	2.26	2.89	3.69	4.75	6.18

Section 6 shows a lognormal maximum likelihood estimate of the gross wage rate distribution for 1983. The maximum likelihood estimates for μ and σ are 7.921 and 0.534, respectively. Hence, Skewness = 1.913.

Gamma distribution: Skewness = 2·(Standard Deviation/Mean).
Estimating the mean and the standard deviation of the gamma distribution by the mean and the standard deviation of the gross wage rate data, we obtain Mean = 3.18 and Standard Deviation = 1.98 (in £); and therefore Skewness = 1.245.

The next two distributions are special cases of the gamma distribution.

Exponential distribution: Skewness = 2, irrespective of the value of its parameter.

Chi-square distribution: The χ^2_n-distribution depends only on the number of degrees of freedom n and converges for $n \to \infty$ to the symmetric normal distribution. We have Skewness = $2\sqrt{2}/\sqrt{n}$. For example, if n = 6833 (the size of the 1983 FES sample of "all workers"), we obtain Skewness = 0.034.

5) The first three initial moments of the kernel estimator are given by:

$$m_1 = \int x\hat{\rho}(x)dx = \frac{1}{n}\cdot\Sigma\, x_i,$$

$$m_2 = \int x^2\hat{\rho}(x)dx = h^2\cdot\int y^2 K(y)dy \; + \frac{1}{n}\cdot\Sigma\, x_i^2$$

and

$$m_3 = \int x^3\hat{\rho}(x)dx = h^3\cdot\int y^3 K(y)dy + 3\cdot h^2\cdot(\frac{1}{n}\cdot\Sigma\, x_i)\cdot\int y^2 K(y)dy + \frac{1}{n}\cdot\Sigma\, x_i^3.$$

The third central moment of $\hat{\rho}$ is given by $m_3-3m_2m_1+2(m_1)^3$; the variance of $\hat{\rho}$ is given by $m_2-(m_1)^2$. The variance of the gross wage rate data for 1983 is $(1,980)^2$; the variance of the standard normal kernel is 1; and h=250 in Figure 2.

6) This assumption is not restrictive. If the unknown density function ρ has a larger range than [a,b], we actually estimate the truncated density

$$\tilde{\rho}(x) = \begin{cases} \alpha\rho(x), & a \le x \le b \\ 0 & , \text{otherwise}, \end{cases}$$

where $\alpha = (\int_a^b \rho(x)dx)^{-1}$.

7) There is a large literature on the female-male wage gap and on the broader issue of discrimination (see, e.g., Becker, 1985; Blinder, 1973; Corcoran and Duncan, 1979; and Duncan and Hoffman, 1978). An interesting recent review of female-male wage differentials and policy responses is given by Gunderson (1989). The survey focuses mainly on studies which pertain to the United States, where most of the empirical work has been conducted. A valuable source for the United Kingdom is Zabalza and Tsannatos (1985). The standard procedure to analyse the determinants of the female-male earnings gap is to estimate earnings equations for samples of men and women separately by the method of least-squares. One then compares the estimated coefficients with each other.

8) See, e.g., Blinder (1980) for the United States; and Gösecke and Bedau (1974) for the Federal Republic of Germany.

9) In a recent article Gomulka and Stern (1990) explore the determinants of the increase in the proportion of employed married women in the United Kingdom over the period from 1970 to 1983. The authors ask how much of the rise can be attributed to changes in the coefficients of their regression model as opposed to changes in the explanatory variables. The data set used is the FES. The measure of employment is a (0-1) variable which describes whether or not a women is doing paid work. Explanatory variables for employment include age, wage and husband's income. The authors provide a detailed description of household structure. Finally, there are regional dummies. The hypothesis of no change in the coefficients over time is strongly rejected, and the authors conclude that "the changing coefficients constituted a major element in the explanation of the rise in the proportion working, accounting for around 65-75 per cent of the change, i.e. 6-8 percentage points out of 9-10". They found that "practically no change was associated with the coefficients on family structure, and among the changing coefficients, the effects of regions appear to be most dominant with an apparent tendency for non-metropolitan regions to become more like London and the South-East in the propensity for wifes to work". Turning to changes in explanatory variables, Gomulka and Stern write that "the decline in the number of children seems to account for around 4 of the overall 9-10-percentage-point increase in the proportion of women working". (The quotations are taken from the last page of the article.)

10) For a discussion of measuring inequality and ranking distributions see, e.g., Atkinson (1970) and Shorrocks (1983). The two articles pay attention to the normative judgements (i.e., the concepts of social welfare) underlying any evaluation of alternative allocations of resources and derive operational rules for empirical work. Both authors refer to the distribution of income. However, the concepts they explore can be applied to distributions of wage rates as well. (Atkinson explicitly remarks that he refers only for convenience to income.)

11) In exploring poverty, researchers typically focus on household income as a measure for standards of living; Atkinson (1987) discusses three basic issues in measuring poverty, namely the choice of the poverty line and that of the poverty index, and the relation between poverty and inequality. When looking at poverty within the work force (and hence turning away from the family to the individual) the analysis should be based on the hourly wage of an individual and not on his labour income. A worker who might be considered as "poor" when looking at his hourly wage can, in principle, compensate this low wage rate by working many hours per week. Accordingly, his labour income does not have to be below a specified poverty standard (e.g., the Supplementary Benefit level in Great Britain). However, whether or not a worker has to be considered as "poor" should not depend on the number of hours worked, but on his or her position in the range of wage rates offered by society.

12) Apart from the impressive work of K. Hildenbrand and W. Hildenbrand (1986), notable exceptions are Deaton (1988) and Härdle and Jerison (1988).

13) Lydall suggested his model first in (1959); at that time essentially the same model had already been put forward by Simon (1957). The work of Lydall and Simon stimulated some interest in the study of hierarchically organised firms. Since the analysis of business firms which are pyramidal in form may lead to a better understanding of the distribution of wages in a market economy, it is interesting to comment briefly on what has been done in the literature.
The optimum size of such firms was investigated (see Williamson, 1967; Beckmann, 1977; and Calvo and Wellisz, 1978) and endogenous explanations of the internal wage and labour utilisation structure were proposed (see Calvo and Wellisz, 1978, 1979; Malcomson, 1984; and Szymanski, 1987). A general feature of these models is that wage differentials across the layers of a hierarchy are inexplicable in terms of differences in labour quality and difficulty of tasks. For instance, Malcomson and Szymanski show that wage

rates tend to rise more over the employment cycle than productivity does (this feature of the labour market has been documented, for instance, in the empirical study by Medoff and Abraham, 1980). Furthermore, it was possible to establish that in a production economy with hierarchically organised firms the wage distribution is skewed to the right relative to the underlying ability distribution (see Rosen, 1982; and Waldman, 1984). The articles by Rosen and Waldman are attempts to sketch a theory of the joint distribution of firm size and earnings generated by market assignments of personnel to hierarchical positions in firms. Both articles are in the spirit of Tuck (1954). Tuck gives an example showing how a hierarchically organised industry may generate a lognormal earnings distribution, although all individuals have identical abilities at the beginning of their work life; see also Mayer (1960), who shows how a normal distribution of ability may lead to a lognormal distribution of earnings. Two further interesting papers are Stiglitz (1975) and Mirrlees (1976); Stiglitz is more generally concerned with the demand for supervision and the return to having a hierarchical production structure, while Mirrlees also tries to explain the distribution of incomes within a firm.
The results of the above authors are interesting, but there are important issues still to be settled. In fact, it even may be impossible to explain the distribution of earnings in a market economy on the basis of purely economic arguments. This view is taken by Simon (1957), who concludes his paper by writing that "it would appear that the distribution of executive salaries is not unambiguously determined by economic forces, but is subject to modification through social processes that determine the relevant norms" (these norms are: first, a norm for the "steepness" of the organisational hierarchies; second, a norm for the wage differential between the supervisor and his subordinates). It would be interesting to link Simon's sociological explanation of wage differentials with Thurow's (1976) "job structure theory", where wages are fixed by employers and an unemployed cannot bid back into his old job at a lower wage.

14) The standard goodness-of-fit test is the χ^2-test developed by Karl Pearson in 1900. For instance, Muellbauer remarks in Atkinson (1976) that "chi-square tests almost always reject two-parameter forms of income distributors; in fact, not even Champernowne's (1952) three- or four-parameter distribution works terribly well" (p. 93). We have not used the χ^2-test here for the following reasons. Firstly, we did not want to divide the range of the unknown distribution into intervals. Secondly, when the sample is large, the test often detects even small departures from the null hypothesis (see, e.g., Cochran, 1952, p. 335); Muellbauer, however, seems to prefer the χ^2-test to the test which we use in Section 6 (see Atkinson, 1976, p. 93). The reader may find the following two contributions interesting:
Harrison (1981) re-examined the, as he calls it, "conventional wisdom" on the size distribution of earnings, using data from the British New Earnings Survey for the year 1972 and two different test statistics: the χ^2-test and a goodness-of-fit test proposed by Gastwirth and Smith (see the reference given by the author). The hypothesis tested was that the earnings distribution is lognormal, but having an upper tail which is better described by the Pareto distribution. Harrison first tested the null hypothesis for the entire population of full-time male workers aged 21 and over; the variable chosen was "gross weekly earnings". He then disaggregated the sample into 16 occupational groups and tested the null hypothesis for each of these groups. He found that "if standard levels of significance are used in χ^2 tests, the conventional wisdom is not supported by the observed distribution of earnings among all workers. The distribution is not lognormal, nor is there evidence of a Pareto upper tail among the top 15-20% of the workers. If the results of the Gastwirth-Smith criterion are considered instead, however, both hypotheses find some support...Similar remarks can be made of the results from distributions within most occupational groups" (p. 628). Turning to the model of hierarchical earnings proposed by Lydall, Harrison concludes in his summary "that Lydall's model cannot easily be advanced as an explanation of the stable Pareto upper tail in the overall

distribution" (p. 630). Notice that the existence of a Pareto upper tail is only supported by the data if we trust - as the author obviously does - the Gastwirth-Smith criterion more than the χ^2-test.
Using the FES data, K. Hildenbrand and S. Islam (1985) compared lognormal, beta and gamma maximum likelihood estimates of the normalised distribution of household income with a nonparametric DMPL-estimate (see Subsection 3.2) and the empirical cumulative distribution function of the normalised income data. They estimated the income density for each odd numbered year from 1969 to 1981 on the sample of "all households". Since the DMPL-estimates turned out to be bimodal (see Section 5), none of the three parametric models describes the income distribution very well. The authors infer from the comparisons that "the gamma distribution fits the sample best" (p. 6). The lognormal and the beta density functions turned out to be very similar.

15) The Kolmogorov D-test is the most widely known goodness-of-fit test based on the empirical distribution function. A detailed discussion of the test can be found, for example, in Kendall and Stuart (1973, pp. 468-478) and Darling (1957); Stephens (1974) is a practical guide to the use of tests based on the empirical distribution function. Because of the strong convergence of the empirical to the true distribution function, the D-test is consistent. A drawback of the test is that it is not unbiased; an example in which it is biased was given by Massey (1950, 1952) who also gave a lower bound for the test's power in large samples.

16) We have in mind the Cramér-Smirnov-von Mises test (see, e.g., Kendall and Stuart, 1973, pp. 466-468; or Darling, 1957).

Chapter 3

An Empirical Investigation of the Labour Market
Part II: Labour Supply and Net Earnings Functions

1. Introduction

In this chapter we continue our analysis of the 1970-85 FES. We now turn to the shape of labour supply and net earnings functions. The statistical labour supply function assigns to any given wage rate w the average value of the variable "weekly hours of work" in the subgroup of individuals receiving the wage rate w; in the case of the net earnings function the response variable is "net weekly earnings". We will occasionally use the term "labour supply schedule"; the terms "function" and "curve" will be used interchangeably.

There is a large body of empirical work on labour supply. The literature has essentially focused on estimating the parameters of certain functional forms for the labour supply function. Whether or not strong distributional assumptions are required depends upon the questions one wants to study. Presumably, many applied economists would agree with Schultz (1980, p. 25): "to estimate the parameters of labor supply responses that could be useful for policymakers, a number of relatively strong assumptions are needed". However, in Chapter 3 we merely want to explore whether the labour market fulfills the "law of supply" which says that a rise in the wage level leads to an increase in aggregate labour supply. To answer this question, we have to estimate the unknown labour supply curve as accurately as possible. As in Chapter 2, we will pursue a nonparametric approach.

The model underlying our study (Section 2) is a generalisation of that explored in Chapter 1. We now consider a population of individuals who do not only differ with respect to their wage rate but also with respect to exogenously given personal characteristics. In general, there is no relationship between individual and aggregate labour supply. The shape of the latter function will depend upon the joint distribution of wage rates and personal characteristics. Microeconomic theory does not suggest any specific form for the statistical labour supply function; one may still agree with Dierker (1974, p. 116): "The question which measures give a good description of a consumption sector typical for a modern economy has not been studied". (See also Chapter 0.)

We do not model family decision-making; we also do not attempt to explain unemployment. There are two types of individuals in the economy about which we have no information. Firstly, there are persons who would like to work but do not find a job. Secondly, some individuals are not in paid employment because their market wage does not exceed their reservation wage (e.g., women having young children). However, we do not know the market wages nonworkers could earn.[1] Strictly speaking, we do not estimate the aggregate labour supply function of a given group of individuals, but merely the relation between wage rate and labour supply of those persons already at work. In other words, our estimates suffer from a "sample selection bias".

There are two further points which we should mention. Firstly, the current market wage of an individual reflects in general past labour supply decisions and hence is not an exogenous variable such as the price of a consumer good. Secondly, our wage variable is definitionally related to the dependent variable; we have

observed wage rate = normal gross earnings/normal weekly hours.

This implies a spurious correlation between explanatory and dependent variable: the measurement error in the wage rate is correlated with that in hours of work. Thus, there may be a bias in the estimates due to "endogeneity of the explanatory variable" and "errors-in-variables". This should not be forgotten when looking at the diagrams plotted in Sections 3-5.

These three sections are organised as the corresponding sections of Chapter 2. Section 3 introduces the methods and applies them to the 1983 FES sample of "all workers"; Section 4 is concerned with subsamples of this data set. In Section 5 we will see that the shape of the labour supply function for the entire population of workers is remarkably stable in the years 1971-85; the function values, however, decrease somewhat.

In Section 6 we compute the elasticity of labour supply (resp. net earnings) with respect to the wage level. We will see that observed labour supply is not very responsive to small variations in the wage rate. Loosely speaking, our estimations confirm "empirical laws of labour supply" derived from parametric micro-econometric models. Section 7 compares our work with the literature. Studies which have used the FES data are, for example, Atkinson and Stern (1981), Blundell and Walker (1982, 1986), K. Hildenbrand and W. Hildenbrand (1986) and Härdle and Jerison (1988).

K. Hildenbrand and W. Hildenbrand (1986) applied nonparametric smoothing methods to the FES expenditure and income data in order to investigate whether the commodity markets fulfill the famous "law of demand". This chapter builds on their ideas; Section 5 is closely related to Härdle and Jerison (1988) who study cross-section Engel curves over time. We remark that our work is not related to the nonparametric consumer analysis of Varian (1982, 1983). In Section 8 we make some concluding commnents.

2. Theoretical Framework

Our simple static labour supply model is set up as follows. There are n consumer goods in the economy supplied by firms to the households, while individuals supply various types of labour to the firms. We assume that the firms offer only linear wage schedules; furthermore, there is a tax on labour income. Let $t: R_+ \longrightarrow R$ denote the tax function. An individual with gross earnings y has to pay t(y) to the government in case of $t(y)>0$; if $t(y)<0$, then the individual receives the amount $-t(y)$ from the government.

Given the commodity price system $p \varepsilon R^n_+$, the gross wage rate w and the tax function t, an individual of type i, $i \varepsilon I$, chooses a consumption plan $f^i = f^i(p,w,t) \; \varepsilon \; R^n_+$ and decides to supply $l^i = l^i(p,w,t)$ hours of a specific type of labour per period (note that l^i may equal zero). The parameter i is used to take account of observable and unobservable personal characteristics (such as age, sex and non-labour income; more generally: tastes for work and preferences for consumer goods) which are exogenous to the consumption and labour supply decision; the set I is called the *set of types.* We denote the *net labour income* of individual $i \in I$ at (p,w,t) by $b^i(p,w,t)$, i.e., $b^i(p,w,t) = wl^i - t(wl^i)$.

Our focus of attention is the dependence of labour supply and net earnings on the gross wage rate. We consider a population of individuals who all face the same consumer goods prices and the same tax function. The individuals differ, however, with respect to their type and the gross wage rate they receive on the labour market. The assumption that p and t are the same for all individuals is made without loss of generality. One can easily make prices and taxes dependent on observable attributes of the particular type $i \in I$ (we have in mind, e.g., price variations across regions). However, this would imply an unnecessary complication of the notation.

Let μ denote a joint distribution of types and gross wage rates (i.e., μ is a probability measure on $I \times R_+$).[2)] Then, for fixed p and t, the per capita labour supply in the population is given by

$$L(p,t) = \int_{I \times R_+} l^i(p,w,t)d\mu(i,w).$$

Notice that the distribution μ is unobservable. Thus, the above definition is not very helpful if one wants to explore the labour market's response to a variation in the wage level. However, the probability measure μ determines a marginal distribution of gross wage rates and for each gross wage rate w a conditional distribution of types i. We consider a "large" labour market. More precisely, we assume that the distribution of gross wage rates can be represented by a (Lebesgue-)density function (i.e., the gross wage rates are continuously distributed on R_+).

Let ρ and $\mu_{|w}$ denote, respectively, the density of the marginal distribution of gross wage rates and the conditional distribution of types given the gross wage rate w. Then the per capita labour supply of individuals with gross wage rate w is given by

$$l(p,w,t) = \int_I l^i(p,w,t)d\mu_{|w},$$

and by a well-known extension of Fubini's theorem (see, e.g., Loève, 1977, pp. 137-138) we can write aggregate labour supply as a function of $l(p,w,t)$ and ρ, i.e.,

$$L(p,t) = \int l(p,w,t)\rho(w)dw.$$

Individuals receiving the gross wage rate w have average net earnings in the amount

$$b(p,w,t) = \int_I b^i(p,w,t)d\mu_{iw}.$$

Consequently, the average value of $b^i(p,w,t)$ in the population can be written as

$$B(p,t) = \int b(p,w,t)\rho(w)dw.$$

In this study we do not explore the dependence of labour supply on the consumer goods prices and the tax function. These are exogenously given throughout our analysis. To shorten the notation, we therefore drop in the following the variables p and t. Thus, $l(w)=l(p,w,t)$ and $b(w)=b(p,w,t)$ for all wage rates w. The functions $l(\cdot)$ and $b(\cdot)$ can be estimated. Speaking in statistical terms, $l(\cdot)$ [resp. $b(\cdot)$] is the regression function for conditional mean labour supply (resp. net earnings) with argument gross wage rate w; the function $l(\cdot)$ is called *labour supply function*, and the function $b(\cdot)$ is called *net earnings function*. We assume that the regression functions are differentiable.[3)] Loosely speaking, we are interested in the average value of the derivatives $\delta_w l(w)$ and $\delta_w b(w)$ with respect to the gross wage rate distribution.

Suppose all gross wage rates will be increased by an absolute amount $a>0$. Then per capita labour supply becomes

$$L(a) = \int l(w+a)\rho(w)dw \qquad (a\varepsilon R_+).$$

In case of a uniform proportional increase of $(\alpha-1)\cdot 100\%$ we have to substitute αw for $w+a$. We set, by abuse of notation,

$$L(\alpha) = \int l(\alpha w)\rho(w)dw \qquad (\alpha \varepsilon R_+).$$

Hence, the derivative $\delta_a L(a)_{|a=0}$ (resp. $\delta_\alpha L(\alpha)_{|\alpha=1}$) gives us the rate of change in per capita labour supply resulting from an absolute (resp. a proportional) increase in all gross wage rates. It follows from the dominated convergence theorem (e.g., Loève, 1977, pp. 126-127) that one may reverse the order of differentiation and integration if (i) $l(\cdot)$ is continuously differentiable, and (ii) ρ is concentrated on a finite interval. Accordingly, we obtain

$$\delta_a L(0) = \int l'(w)\rho(w)dw$$

and

$$\delta_\alpha L(1) = \int l'(w)w\rho(w)dw.$$

Observe that the quotient $\delta_\alpha L(1)/L$ represents the gross wage elasticity of per capita labour supply at α=1.

In the case of net earnings we define analogously:

$$B(a) = \int b(w+a)\rho(w)dw,\ a\varepsilon R_+, \text{ and } B(\alpha) = \int b(\alpha w)\rho(w)dw,\ \alpha\varepsilon R_+.$$

Thus,

$$\delta_a B(0) = \int b'(w)\rho(w)dw \text{ and } \delta_\alpha B(1) = \int b'(w)w\rho(w)dw.$$

For the purpose of the present study, this is all one has to know. Nevertheless, some remarks may be helpful. Let us return to the individual's consumption and labour supply decision. In neoclassical consumer theory the choice (f^i,l^i) is derived from utility maximisation. The set I is usually a subset of the real numbers. To each consumer i in I one assigns a real number m^i, interpreted as the consumer's *non-labour income*, a subset X^i of R^{n+1} and a binary relation $\precsim_i$ defined on X^i; X^i is called the *consumption set* of individual i, and $\precsim_i$ is called his *preference relation*. Given the price system (p,w) - we omit here the tax function -, the individual chooses a point (f^i, l^i) which maximises his preferences subject to the budget constraint $pf \leq wl+m^i$. That is, (f^i, l^i) is feasible for the individual at the price system (p,w) and satisfies $(f^i, l^i) \succsim_i (f,l)$ for all (f,l) in X^i with $pf \leq wl+m^i$ (the standard reference is Debreu, 1959, Ch. 4).

In the present setting we do not presume this kind of rationality. The individual labour supply function is taken as the primitive concept of our simple model. The crucial assumption is that for an investigation of the elasticity of per capita labour supply with respect to the gross wage rate external effects on the labour market can be disregarded. But we have imposed on the functions l^i no other restrictions. Therefore one may think of them as being derived from preference maximisation. It should be emphasised, however, that we do not need neoclassical rationality in order to postulate a functional relationship between (p,w,t) and individual labour supply l^i.

A second point should be mentioned. In many empirical studies labour supply is regressed on the *marginal net wage rate.* In general, the gross wage rate is not considered as a very satisfactory wage measure. For example, Killingsworth (1983) writes in his book on labour supply: "what is in fact required is a measure of the marginal wage rate" (p. 88). The argument behind this statement goes as follows. Consider an individual with convex, continuous and locally non-satiated preferences $\precsim$ defined on his consumption set $X \subseteq R^{n+1}$ (for precise definitions, see, e.g., Debreu, 1959; and Varian, 1984). Let m be his non-labour income, and let $b(l) = wl-t(wl)$.

Then the individual's budget set is given by all consumption and labour supply combinations (f,l) in his consumption set X satisfying $pf \leq b(l)+m$. Let (f^*,l^*) maximise the individual's preferences over this set. By local non-satiation, we have $pf^*=b(l^*)+m$. Because of the tax function t the budget equation $pf=b(l)+m$ is typically non-linear. However, if we set $w' = b'(l^*) = [1-t'(wl^*)]w$ and $m' = pf^*-w'l^*$, then the hyperplane

$$H(p,w',m') = \{(x,l) \in R^n \times R: px = w'l + m'\}$$

separates the point (f^*,l^*) from the (convex) set of consumption and labour supply combinations which are preferred to (f^*,l^*) by the consumer.

This means that if a point (f,l) is preferred to (f*,l*), then it costs with respect to the price system (p,w') at least as much as (f*,l*). Since preferences are continuous, one easily verifies that there exists no commodity bundle (f,l) which is strictly preferred to (f*,l*) but does not cost more. Consequently, the consumer will choose (f*,l*) regardless of whether he is faced with the budget set belonging to (p,w,t) or the "linearised" budget set obtained by, technically speaking, intersecting the closed half-space below H(p,w',m') with the consumption set X.[4]

The marginal net wage rate w' and the non-labour income m' are also called "linearised values" of w and m. Figure 1 illustrates the situation for the case n=1.

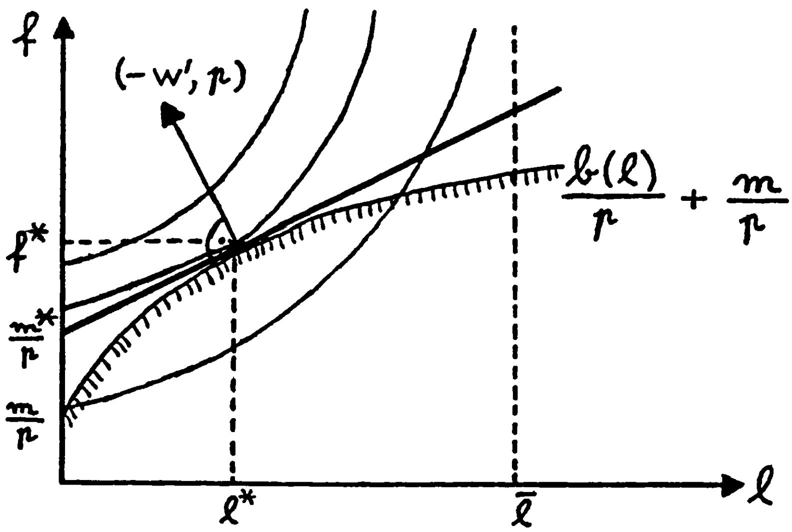


Figure 1

The above relation between the two budget sets shows "that cases with complicated budget constraints can be simplified by "linearization" - that is, by converting such constraints into their straight line equivalent - at each individual's equilibrium point" (Killingsworth, 1983, p. 90). It is then concluded in the literature that the marginal net wage rate w'=(1-t')w is the relevant wage variable on which labour supply should be regressed. Of course, this is true in the neoclassical model if one compares the marginal net wage with the average hourly net wage as a possible explanatory variable of labour supply.

However, the simple "story" behind our model is that employers fix the gross wage rates w, the invisible Walrasian auctioneer calls a given set of prices p, and the government imposes a tax on earned income. Wage rates, consumer goods prices and the tax function are exogenous to labour supply. Consequently, individual labour supply will, amongst other things, depend upon them. Our concern, in turn, is to find out how a particular change in the gross wage rate distribution will affect the average value of the individual labour supplies. Hence, in this set-up we do not need the marginal net wage rates.

In addition, from the point of view of this study, we do not know whether individuals solve their labour supply decision problem rationally. It therefore seems to us safer to express labour supply as a function of the gross wage rate.

3. Nonparametric Regression Curve Estimation

Let $\{(y_i, w_i): i=1,\dots,n\}$ be a random sample from a bivariate distribution with joint density $\rho(y,w)$. Let $\rho(w)$ denote the marginal density of w, and let $\rho(y|w)$ denote the conditional density of y given w. Our aim is to estimate the average value of y for given w, i.e., the regression function

$$g(w) = E(y|w) = \int y\,\rho(y|w)\,dy.$$

Recall that $\rho(y|w) = \rho(y,w)/\rho(w)$, provided $\rho(w)>0$. Thus, we can proceed by either estimating $\rho(w)$ and $\rho(y,w)$ from the sample and then computing $g(\cdot)$, or by directly estimating the regression function. In this study we will directly estimate $g(\cdot)$ without making a priori assumptions about its functional form.

A description of the data was given in Chapter 2. In our case, $\{(y_i, w_i): i=1,\dots,n\}$ is a sample of workers taken from the FES; w_i denotes the gross wage rate of worker i and y_i stands for his or her weekly hours of work (resp. net earnings). Hence, the graph of $g(\cdot)$ is the statistical labour supply (resp. net earnings) curve of the population from which the sample was drawn. An estimator of the unknown regression function $g(\cdot)$ will be denoted by $\hat{g}(\cdot)$.

In Chapter 2 we have discussed in more detail the ideas underlying density estimation. Since nonparametric regression is essentially a variant of the same theme, we will be briefer here. Excellent discussions of the techniques introduced below (and of related methods) can be found, e.g., in Eubank (1988), Gasser and Müller (1979), Härdle (1990), Prakasa Rao (1983) and Stone (1977).

In all cases the unknown regression function was estimated over an interval $[0,w_{0.99}]$, where $w_{0.99}$ denotes the 99th percentile of the empirical distribution of gross wage rates; the corresponding density estimates are shown in Chapter 2. The diagrams relating to this section are plotted on pages 176-180; see Figures 1, 2 and 5 of Chapter 2 for the density functions (pp. 86-92). Recall that the FES earnings data are recorded in tenths of pence.

3.1. Naive Estimation and Spline Smoothing

By definition, the regression function $g(\cdot)$ assigns to each value of w the mean of the corresponding y-values. One can therefore construct a "naive estimate" of the labour supply function as follows: Fix an interval $[a,b]$ that contains the gross wage rate data; select a grid of points $a=a_0<a_1\ldots<a_m=b$, and put $A_i = [a_i,a_{i+1}[$ $(i=0,\ldots,m-1)$. Let $\bar{l}_i$ denote the mean labour supply in the subgroup of workers whose gross wage rate is in A_i, i.e.,

$$\bar{l}_i = \frac{1}{\#\{j:\ w_j \varepsilon A_i\}}\cdot\Sigma_{j:w_j\varepsilon A_i}\ y_j \qquad (i=0,\ldots,m-1).$$

A natural estimator of the unknown labour supply function is then given by

$$\hat{l}(w) = \begin{cases} \bar{l}_0, & \text{if } w\varepsilon A_0 \\ \cdot & \\ \cdot & \\ \cdot & \\ \bar{l}_{m-1}, & \text{if } w\varepsilon A_{m-1}. \end{cases}$$

Notice that the definition of the estimator is very similar to that of the histogram (see Chapter 2, Section 3); Tukey calls $\hat{l}(\cdot)$ therefore the "regressogram". Obviously, the shape of the curve will depend on our choice of the partition $(A_i)_{i=0,\ldots,m-1}$. Figure 2a shows a naive estimate of the labour supply function for the population of "all workers 1983". The estimate was constructed using an equally spaced mesh of points $a_0,\ldots,a_m$ with $a_0=100$, $a_m=10000$ and $m=40$. As we see, the estimate becomes extremely unstable in the upper range of the gross wage rate distribution. We observe, however, that mean labour supply exhibits the tendency to decrease for large values of w.

The gross wage rates are not uniformly distributed over the interval [100,10000] and this should be taken into account when choosing the interval partition. We remark that around 75 (resp. 90) per cent of the workers earn per hour not more than £3.8 (resp. £5.4). Recall that the size of the FES sample is n=6833; 6764 individuals have a gross wage rate rate between £0.1 and £10. In the range from £6.6 to £10 we have 338 workers; 68 individuals reported normal earnings and normal hours of work implying a gross wage rate lower than £0.72.

In order to prevent rapid fluctuations of the estimator we have to increase the intervals A_i in the upper tail of the distribution. Figure 2a also suggests that on should increase A_0 considerably. This was done in Figure 2b. The empirical labour supply function is now "strictly" decreasing in the upper range of the gross wage rate distribution. In the second diagram of Figure 2b the points $(\bar{l}_j,t_j)_{j=0,\ldots,m-1}$, where $t_0=0$ and $t_j = (a_j+a_{j+1})/2$ $(j=1,\ldots,m-1)$, were interpolated by a cubic spline (see Subsection 3.2 of Chapter 2).

Figure 2b shows that the spline interpolant is not a very good smooth approximation to the regressogram. One sees immediately that in the range from 0 to £2 a cubic polynomial was used, and in this range the deviation between the two functions is also large. Furthermore, the spline interpolant exhibits local fluctuations which we would like to avoid. One can obtain a better smooth approximation to $\hat{l}(\cdot)$ by varying the points through which the interpolant passes. However, there exist more sophisticated smoothing techniques. One such method is kernel smoothing.[5] Generally speaking, kernel smoothers and related methods can be thought of as "stabilised" naive estimators.

3.2. Kernel Estimators

Let K: R⟶R be a kernel function, i.e., K is a probability density with $K(w) = K(-w)$ for all $w \in R$. Let $h>0$, and let

$$m_i(w) = \frac{K(\frac{w-w_i}{h})}{\Sigma_{j=1}^{n} K(\frac{w-w_j}{h})}, \qquad w \in R \quad (i=1,\ldots,n).$$

Then the function g defined by

$$(K) \qquad g(w) = \Sigma y_i m_i(w), \qquad w \in R,$$

is called *(ordinary) kernel regression estimator* with *kernel* K and *smoothing parameter* h.

Notice that $m_i(w) = \frac{1}{nh} \cdot K(\frac{w-w_i}{h}) / \hat{\rho}(w)$, where $\hat{\rho}$ denotes the kernel density estimator defined in Chapter 2. Thus, the regression estimation yields as a by-product an estimate of the unknown wage rate density ρ. The function $m_i(\cdot)$ is a weight function assigned to the observation (y_i, w_i). Because of the denominator in the definition of the $m_i(\cdot)$, the weights sum up to one. Hence, (K) is simply a weighted average of the observations $y_1, \ldots, y_n$. The weights are more equally distributed for large values of h than for small values; the larger h, the more individuals will be included in the average (K) and, consequently, the smoother will be the regression.[6] As in the case of density estimation, the crucial condition for pointwise consistency of the kernel estimator is that h has to converge to zero as the sample size goes to infinity, but not as rapidly as n^{-1}.

If the regressor w is not uniformly distributed, then one can improve the estimation by varying h across the sample. We would like to choose h small (resp. large) in regions where the w_i lie very dense (resp. where we have only relatively few observation). Let $\hat{\rho}$ be the ordinary kernel density estimator with kernel K and smoothing parameter h. Let μ denote the geometric mean of the function values $\hat{f}(w_i)$, $i=1,\ldots,n$. Let $t=(t_1,\ldots,t_n)$ with $t_i=\{\mu/\hat{f}(w_i)\}^{1/2}$ for $i=1,\ldots,n$. Then the function g defined by

$$\text{(A)} \qquad g(w) = \Sigma y_i m_i(w;t), \quad w\varepsilon R,$$

where

$$m_i(w;t) = \frac{\frac{1}{t_i}\cdot K(\frac{w-w_i}{t_i h})}{\Sigma_{j=1}^{n} \frac{1}{t_j}\cdot K(\frac{w-w_j}{t_j h})}, \qquad w\varepsilon R \quad (i=1,\ldots,n),$$

is called *adaptive kernel regression estimator* with *kernel* K, *local bandwidth factors* $t_1,\ldots,t_n$ and *(global) smoothing parameter* h. (See also Chapter 2, Subsection 3.1.)

Estimators of $g'(\cdot)$ are obtained by differentiating (K) and (A) with respect to w. As in Chapter 2, we chose for K the standard Gaussian density, i.e., $K(w) = \frac{1}{\sqrt{2\pi}}\cdot\exp\{-\frac{1}{2}w^2\}$, $w\varepsilon R$.[7] To find a suitable value for h, we used the formula $h = 1.06\cdot\sigma\cdot n^{-1/5}$, where σ denotes the standard deviation in the sample of gross wage rates. As already mentioned in Chapter 2, this choice for the smoothing parameter copes fairly well with the data. At the very least, it was always a good starting point for subsequent fine tuning.[8]

In Figures 3a and 3b the labour supply function for 1983 was estimated by the ordinary and adaptive kernel method. As we see, per capita labour supply is decreasing on [0,1] (in £), increasing on [1,3.8] (i.e., in the range from around the 3rd to the 75th percentile of f) and decreasing from then onwards; l(w) rapidly increases in the range between £1.2 and £3.0 (the median of f is at w=2.75). The empirical correlation between labour supply and gross wage rate is 0.135.

The adaptive kernel method significantly improves upon the estimation with a fixed window width across the entire sample. Notice that the ordinary kernel estimate with h=250 is extremely noisy in the upper range of f. After having increased the value of h by more than 50 per cent there are still random fluctuations in the regression function for large values of w, while the adaptive kernel method produces an impressively smooth curve. The adaptive kernel method copes also very well with those few individuals who "claimed" that they worked many hours at an almost vanishing gross wage rate. On passing from the ordinary kernel estimation with h=250 to the adaptive kernel estimation with h=380, l(0) drops from around 38 to 28 hours.

Recall that the wage rate is obtained by dividing gross earnings by hours of work. If the informant stated that he worked usually *fourteen* hours and the interviewer wrongly understood *forty*, one obtains a fairly low wage rate. Presumably the bottom, say, 1 per cent of the empirical gross wage rate distribution (= 68 observations) can be explained by such errors and by errors which occurred when the data were recorded on the computer tapes.

Figure 3 shows that individuals in full-time employment receive in general better wage rates than part-time workers (of course, we know that already from Chapter 2). Figure 3 also suggests that full-time workers with high wage rates do not work more hours than those having lower hourly earnings. An adaptive kernel estimate of the labour supply schedule for full-time workers is given in Figure 4. The curve is first decreasing, very gently increasing from around the 10th to the 50th percentile of the gross wage rate distribution for full-time workers (in this range the difference between the maximum and the minimum is less than one hour) and from then onwards, as expected, again decreasing. Excluding part-time workers from the sample of "all workers 1983" leads to an empirical correlation between labour supply and gross wage rate of -0.103 (i.e., the correlation coefficient "changes its sign").

The correlation between net earnings and gross wage rate in the entire sample is 0.787. It is interesting to observe that excluding part-time workers from the sample leads to the larger correlation coefficient of 0.893. Figure 5 shows estimates of the per capita net earnings function for the whole population of workers. The least-squares line plotted there is given by

$$b(w) = 15921 + 20.72 \cdot w, \quad w \in R_+ .$$

For the adaptive kernel smoother the same window width as in Figure 3b was used. As expected from the shape of the aggregate labour supply function, the kernel estimate of b(·) differs considerably from the least-squares line. We see that the graph of the kernel estimate has the same shape as the well-known textbook production function, i.e., first convex and then concave.

Whether or not the slope of the least-squares line is a good approximation to the average slope of the kernel estimate will now depend on how the wage rates are distributed over the interval [0,10000]. Since around 75 (resp. 90) per cent of the gross wage rates are not greater than £3.8 (resp. £5.4), Figure 5 suggests that $\delta_a B(0)$ is substantially higher than 20.72. The precise figures for the years 1970-85 will be given in Section 6.

Notice that one obtains a linear net earnings function if (i) the tax function is of the form $t(y) = \alpha + \beta \cdot y$ and (ii) all individuals work the same number of hours. Clearly, in this case all observations would lie on a straight line. The net earnings function is then given by

$$b(w) = -\alpha + (1-\beta)\cdot\bar{l}\cdot w, \quad w \in R_+,$$

where $\bar{l}$ stands for the number of hours worked.

The arithmetic mean of the labour supply data for the year 1983 is 35.57. Hence, setting $\bar{l}$ = 35.57, we obtain from $(1-\beta)\cdot\bar{l}$ = 20.72 the marginal tax rate β = 0.417. However, the marginal tax rate in Britain is not that high. For the vast majority of full-time workers the marginal tax rate is 34 per cent (see the brief description of the tax schedule on pp. 80-1).

We remark that the shape of the regression functions is robust against reasonable variations of the smoothing parameter. In principle, one could have excluded the bottom, say, 1 per cent from the distribution of gross wage rates in order to avoid giving undue weight to peculiar observations. We have not done this here for two reasons. Firstly, from the point of view of a descriptive data analysis, it is preferable to obtain an impression of the distribution of the data in the entire sample (with all its possible "peculiarities"). Secondly, nonparametric modelling copes very well with peculiarities in the data. In our model the labour supply elasticity is essentially determined by the behaviour of $l(\cdot)$ in the main body of ρ. The behaviour of $l(\cdot)$ for very small and very large wage rates in the range of ρ should not be completely ignored; it will, however, only have a small effect on $\delta_\alpha L(1)/L$.

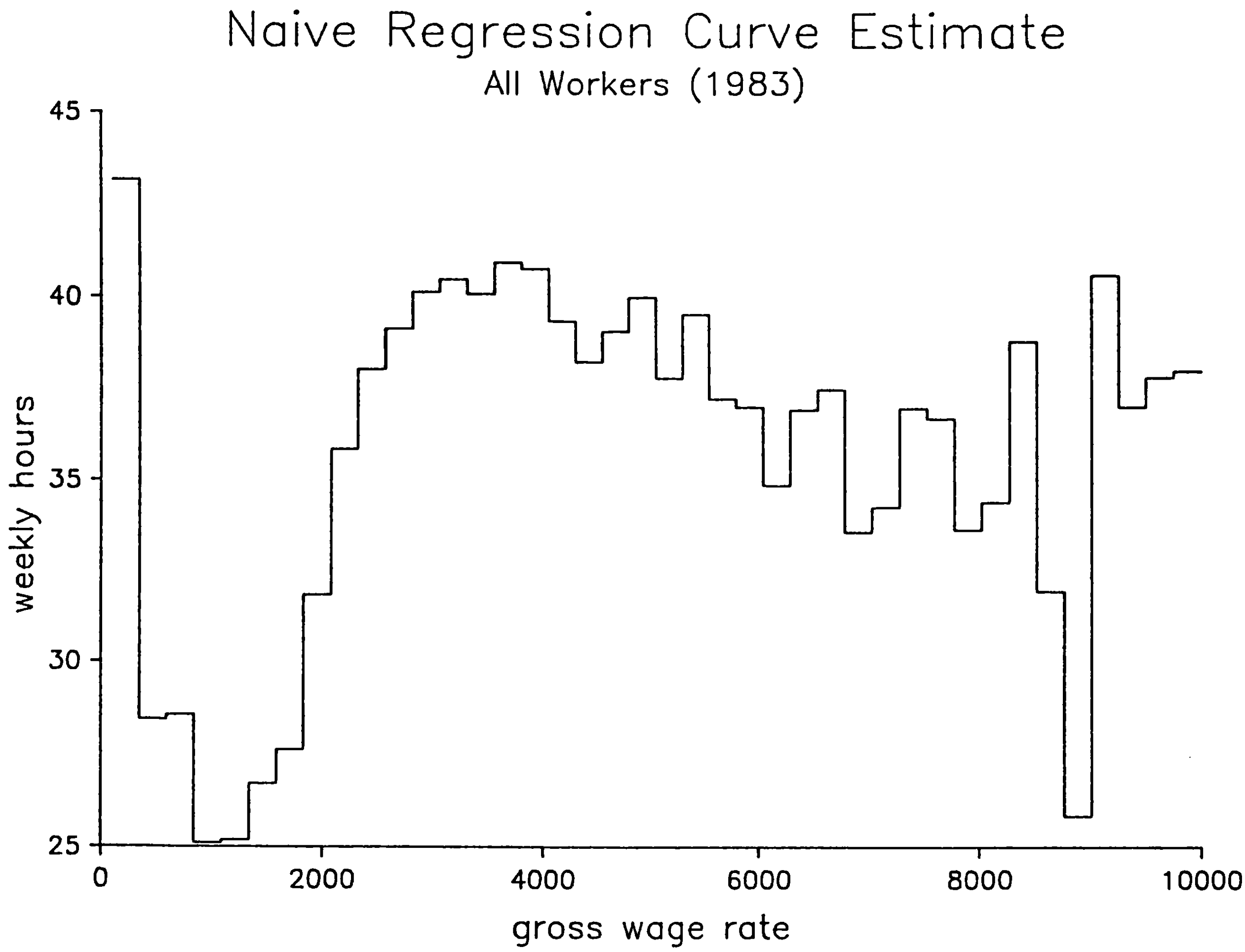
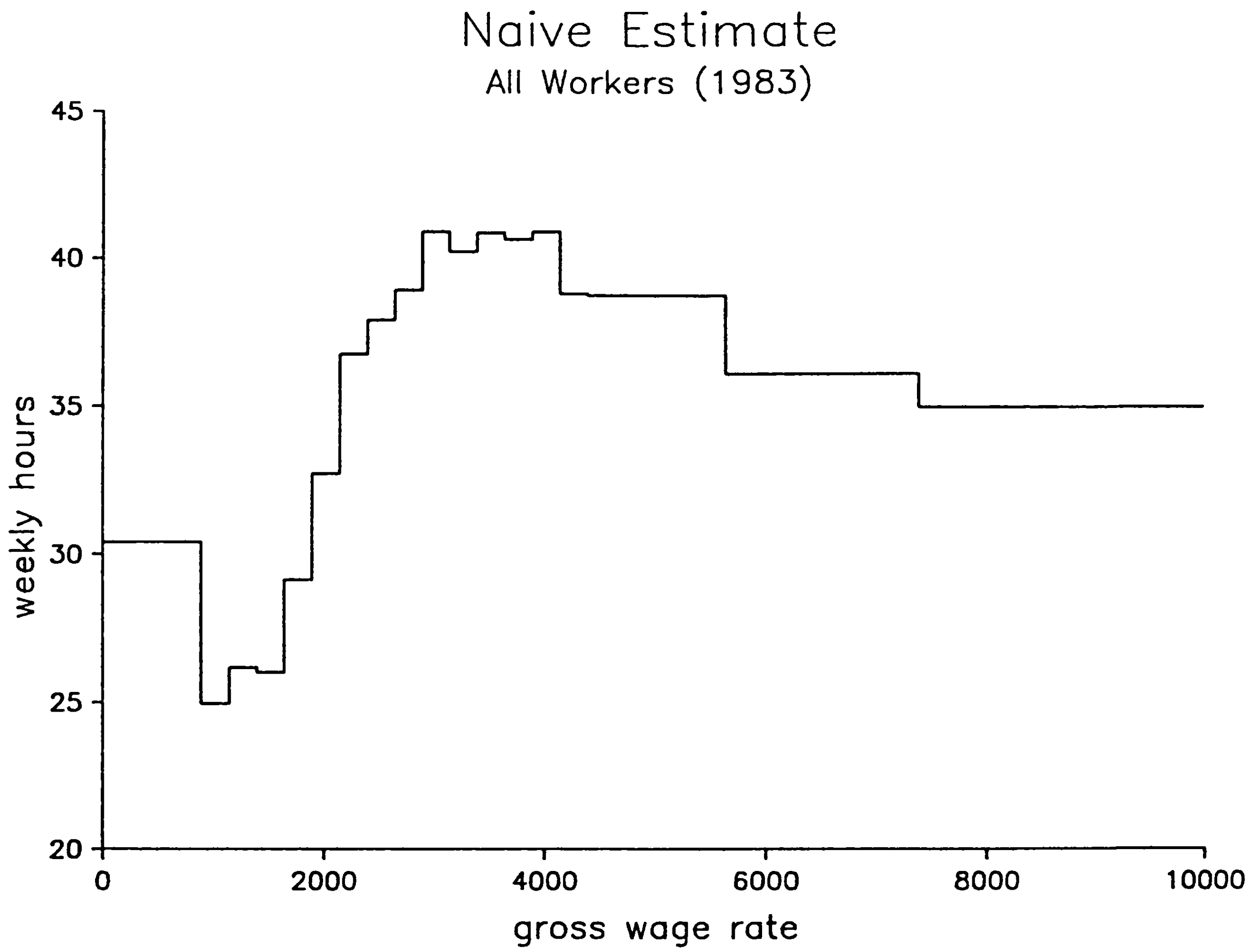


Figure 2a


Naive Estimate
All Workers (1983)
45
40
35
30
25
20
weekly hours
0
2000
4000
6000
8000
10000
gross wage rate

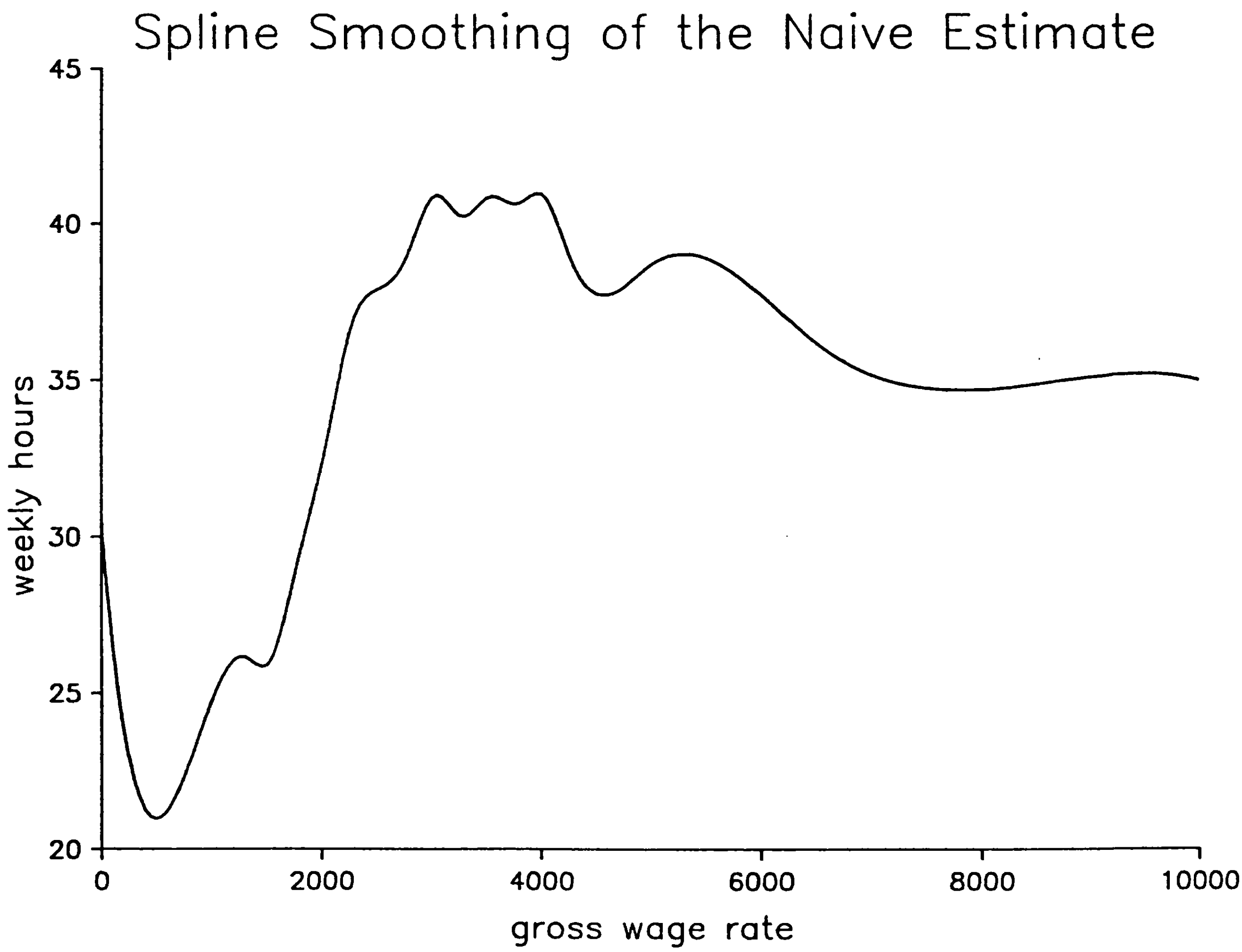

Spline Smoothing of the Naive Estimate
45
40
35
30
25
20
weekly hours
0
2000
4000
6000
8000
10000
gross wage rate

Figure 2b

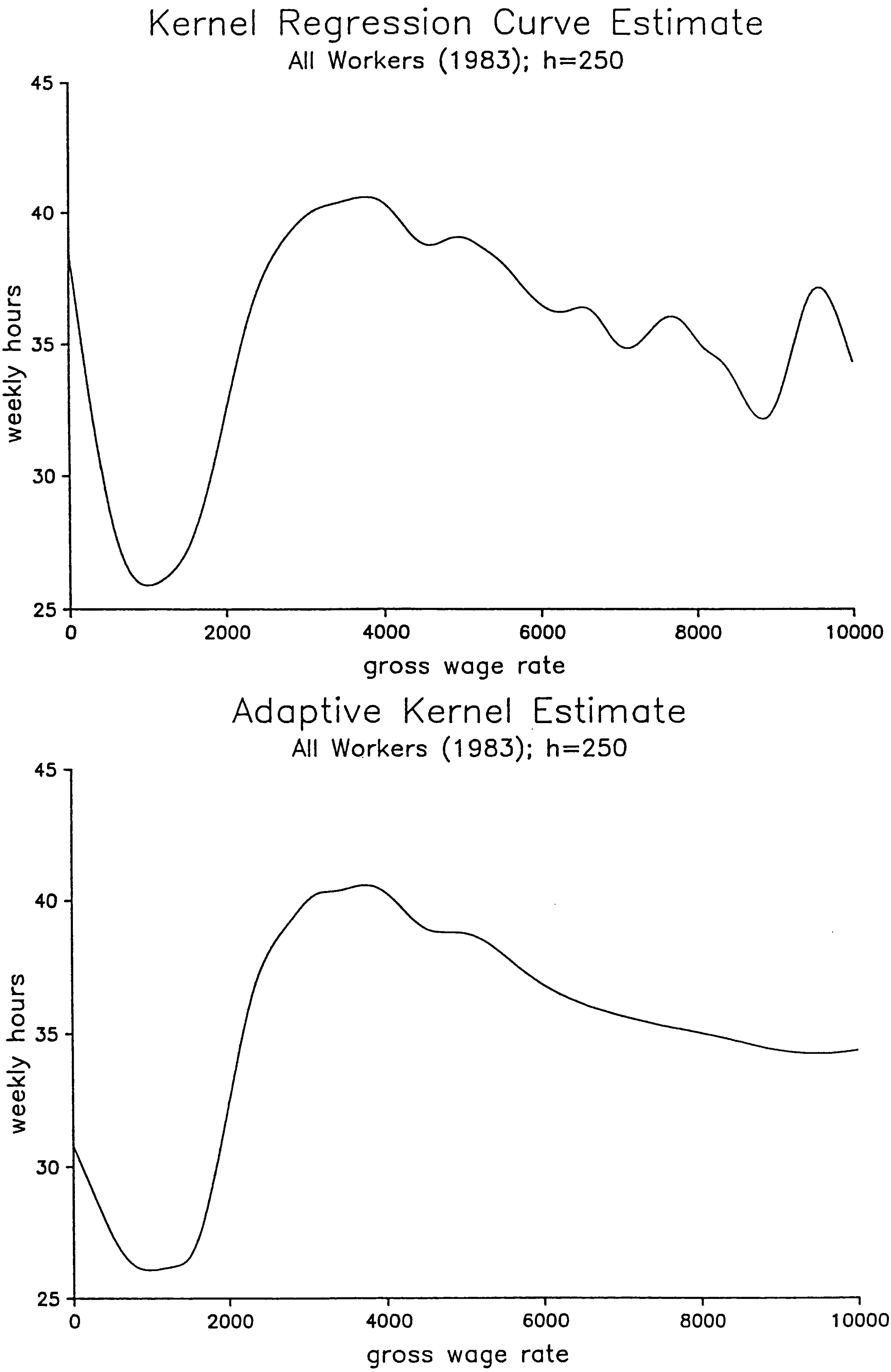
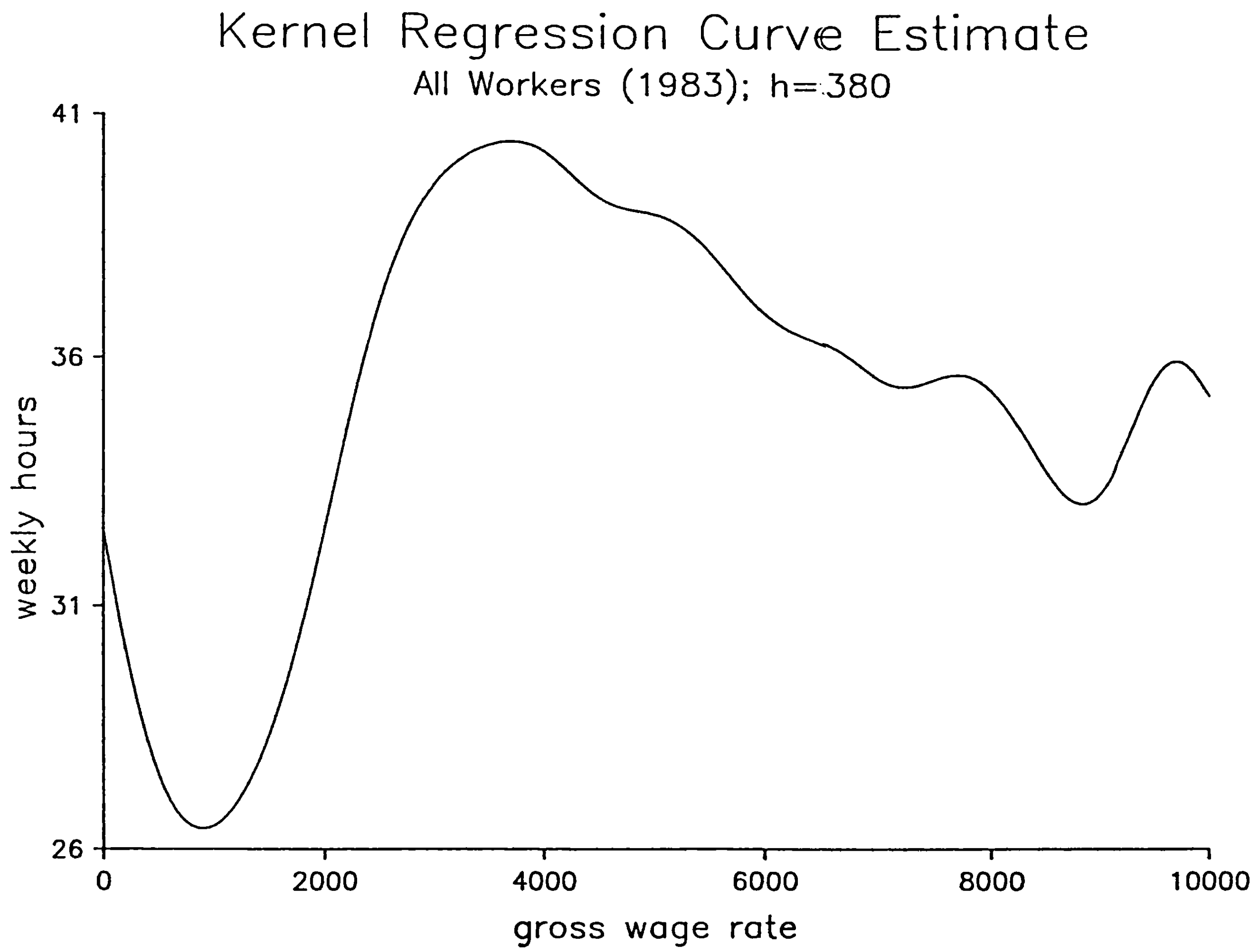

Kernel Regression Curve Estimate
All Workers (1983); h=250
weekly hours
45
40
35
30
25
0
2000
4000
6000
8000
10000
gross wage rate
Adaptive Kernel Estimate
All Workers (1983); h=250
weekly hours
45
40
35
30
25
0
2000
4000
6000
8000
10000
gross wage rate

Figure 3a


Kernel Regression Curve Estimate
All Workers (1983); h=380
41
36
31
26
weekly hours
0
2000
4000
6000
8000
10000
gross wage rate

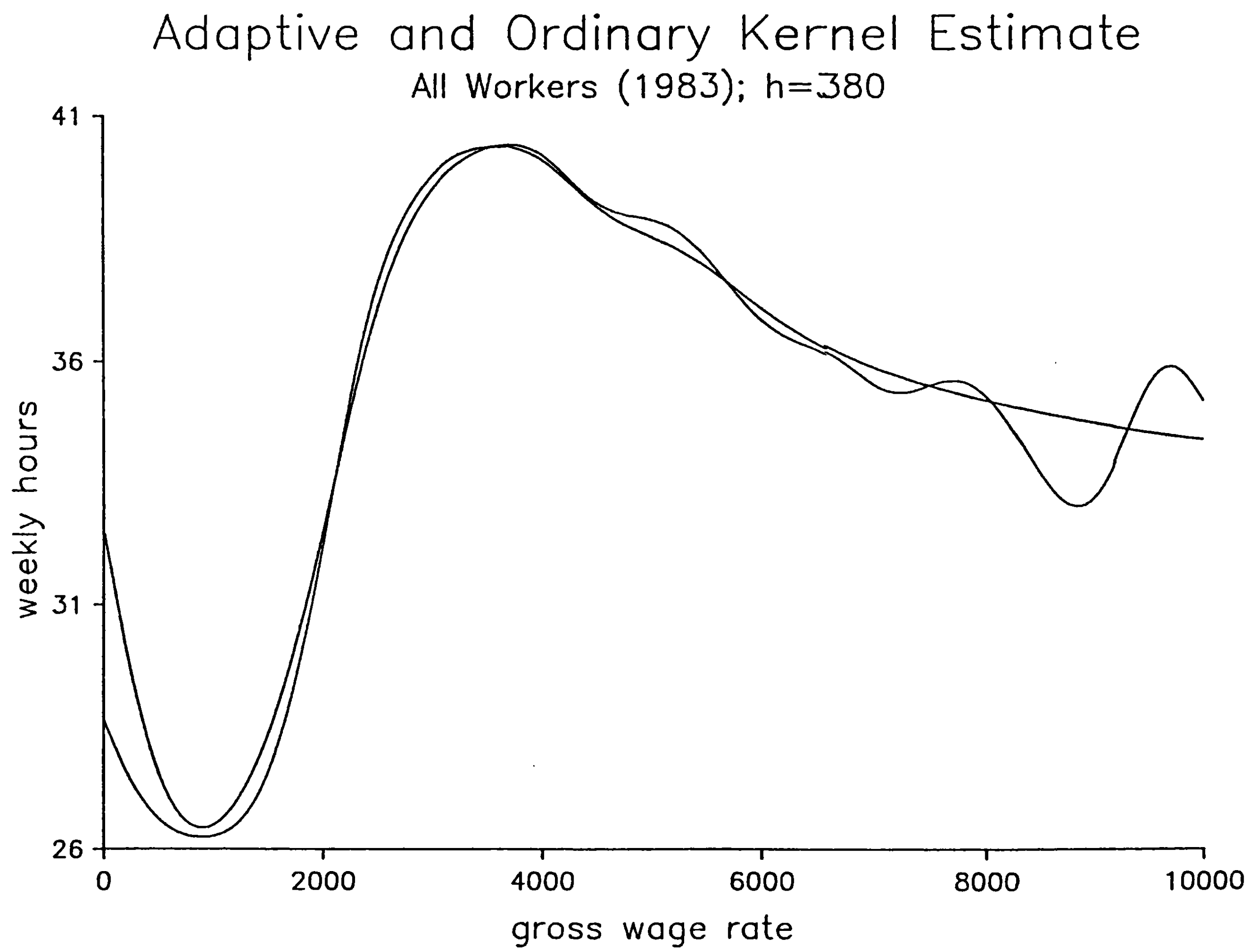

Adaptive and Ordinary Kernel Estimate
All Workers (1983); h=380
41
36
31
26
weekly hours
0
2000
4000
6000
8000
10000
gross wage rate

Figure 3b

Figure 4

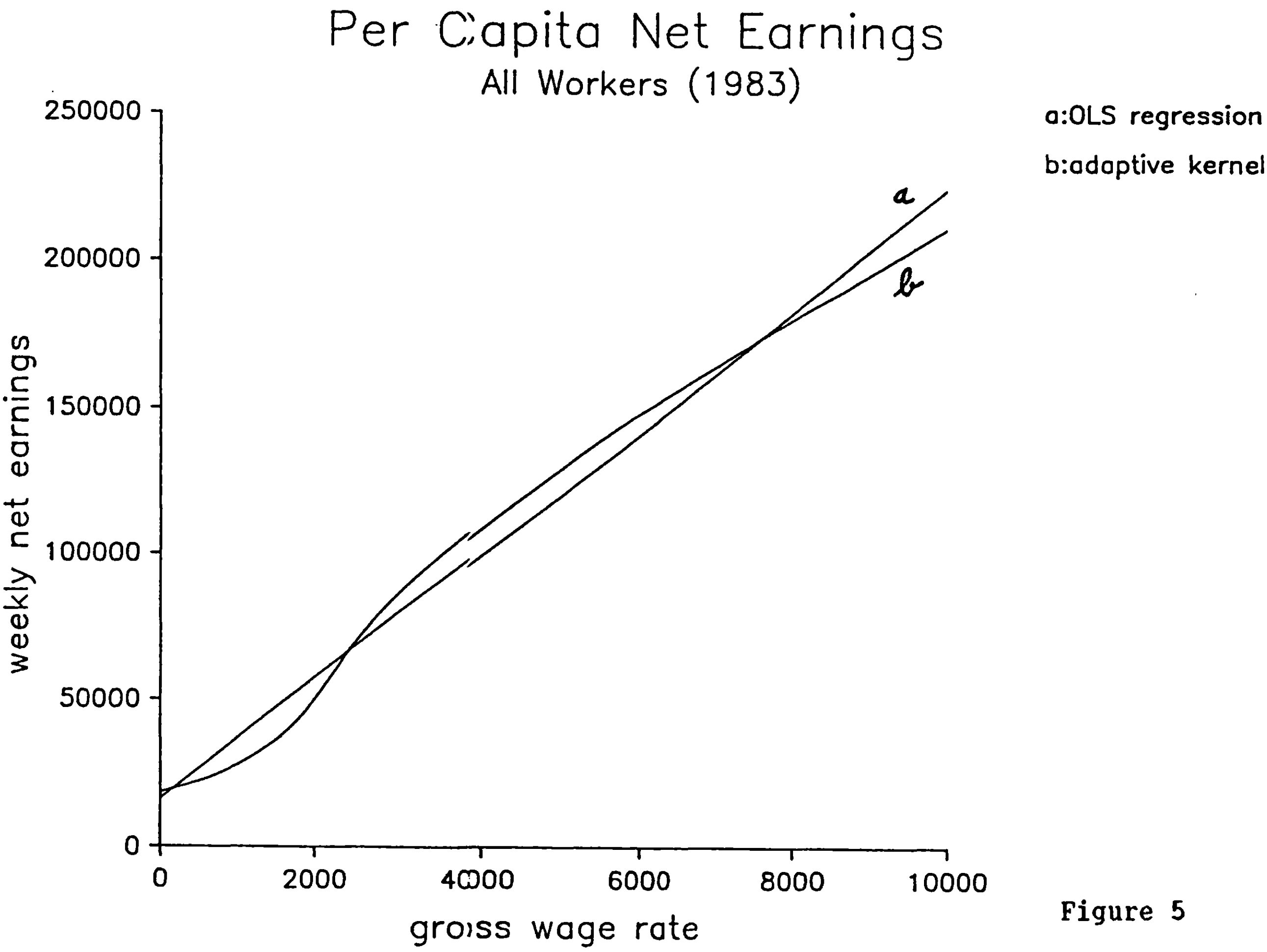


Figure 5

4. Investigating Subsamples

In Section 4 of Chapter 2 we explored for the year 1983 the distribution of the variables "net (resp. gross) weekly earnings", "hours of work" and "gross wage rate" within the following eight populations: female (resp. male) workers, female manual (resp. non-manual) workers, male manual (resp. non-manual) workers, manual (resp. non-manual) workers. Recall that each of these populations includes part-time workers. In this section we continue our analysis of subsamples of the 1983 FES. The section provides estimates for the labour supply and net earnings curves of the above populations; in addition, we will have a look at labour supply schedules for populations of full-time workers. The latter populations are defined by excluding from the above groups those individuals who usually work less than 31 hours per week (of course, this is a very ad hoc definition of full-time employment). The regression curves were estimated on each sample separately. Clearly, the reasons for investigating subsamples are the same as those stated in Chapter 2.

The estimates are plotted on pages 189-196. Figure 6a presents adaptive kernel estimates for the labour supply functions of the eight populations already considered in Chapter 2. Figure 6b compares each of these curves with the corresponding ordinary kernel estimate (i.e., in both estimations the same value for the smoothing parameter h was used). The curves drawn in Figure 7 are adaptive kernel estimates of the labour supply schedules for the populations of full-time workers. Finally, Figure 8 presents estimates of the net earnings functions corresponding to the labour supply functions of Figure 6. In Figure 8a adaptive kernel estimates are shown; in Figure 8b each adaptive kernel estimate is plotted together with the corresponding ordinary kernel estimate. For the net earnings functions the same values of the smoothing parameter were used as for the labour supply functions in Figure 6.

Estimates of the gross wage rate distributions for the populations considered in Figures 6 and 8 are given in Figures 6 and 7 of Chapter 2 (see pp. 102-107). Chapter 2 does not provide estimates of the gross wage rate distributions for the populations of Figure 7. Recall, however, that we mentioned there that excluding part-time workers from the samples does not change the shape of the density functions very much. All wage rate densities are unimodal, i.e., first increasing and then decreasing. Loosely speaking, excluding part-time workers from a sample leads to a density function that is less skewed to the right. Table 1 below provides sample percentiles of the gross wage rate distributions. Table 1a relates to Figures 6 and 8; Table 1b relates to Figure 7.

Table 1a Distribution of Gross Wage Rates
Sample percentiles in tenths of pence; all workers and subgroups, 1983

Sample	1%	5%	10%	25%	50%	75%	90%	95%	99%
All Workers	718	1200	1500	1947	2750	3808	5369	6618	10096
Females	609	1073	1333	1667	2084	2871	4032	5221	7962
Manual Females	497	956	1143	1500	1754	2106	2575	2930	4614
Non-Man. Females	752	1200	1476	1824	2429	3324	4684	6062	8648
Males	833	1418	1853	2500	3267	4416	5997	7376	11167
Manual Males	800	1255	1692	2283	2894	3546	4307	5000	6470
Non-Man. Males	930	1674	2184	3066	4167	5659	7507	9026	13889
Manuals	667	1081	1400	1787	2462	3223	4015	4651	6090
Non-Manuals	807	1316	1615	2141	3118	4599	6387	7732	11869

Table 1b Distribution of Gross Wage Rates
Sample percentiles in tenths of pence; full-time workers, 1983

Sample	1%	5%	10%	25%	50%	75%	90%	95%	99%
All F.-T. Workers	817	1424	1734	2271	3017	4004	5460	6667	10207
Females	564	1184	1481	1848	2380	3132	4055	4848	7008
Manual Females	447	893	1183	1600	1971	2390	2929	3497	4708
Non-Man. Females	708	1324	1633	1983	2599	3391	4424	5123	7424
Males	965	1607	2000	2564	3290	4382	5877	7267	10977
Manual Males	920	1486	1843	2366	2937	3562	4314	5000	6437
Non-Man. Males	1035	1834	2297	3077	4136	5543	7419	8981	13067
Manuals	771	1319	1642	2151	2759	3413	4202	4850	6221
Non-Manuals	1001	1539	1814	2410	3390	4737	6408	7782	11775

Looking at Figure 6b, we see that the ordinary kernel method produces poor results in the lower and upper range of the underlying gross wage rate distribution. In the upper range of ρ the ordinary kernel estimator becomes unstable. In the lower range of ρ the estimator remains "smooth". We see, however, that the ordinary kernel method gives too much weight to "peculiar" observations near w=0.

Comparing the diagrams of this and the previous section with those of Chapter 2, we see that the adaptive kernel method improves the estimation results more in the case of nonparametric regression. In principle, we could have restricted attention to the ordinary kernel method in Chapter 2 (in this case we should have used in all estimations a smaller value for the smoothing parameter in order to avoid obscuring detail in the central part of the densities).

On passing from Figure 6 to Figure 7 the shape of the labour supply functions changes substantially. While the curves of Figure 6 are rapidly increasing before they are eventually decreasing (only the labour supply function for the subgroup "female manual workers" is not decreasing in the upper range of ρ), Figure 7 shows that full-time workers with high hourly earnings typically do not work more hours than those located in the main body and the lower range of the wage rate distribution. Apart from the labour supply schedule for the subgroup "full-time manual workers" the curves plotted in Figure 7 can be approximated by gently downward sloping functions.

This change in the relation between labour supply and gross wage rate can also be seen when looking at the correlation coefficients of the samples given in Table 2 on page 185. In 5 of the 8 samples underlying Figure 6 the correlation coefficient is positive. In the populations of part-time workers labour supply is positively correlated with the wage rate, while we observe for full-time workers a negative correlation between the two variables. There are two exceptions. In the sample of "part-time non-manual female workers" the correlation coefficient has a negative sign and in that of "full-time manual workers" it has a positive sign; both correlation coefficients are approximately zero. The means and standard deviations given in the second part of Table 2 show that the correlation coefficients are fairly stable across the years (see also the tables in the Appendix). In the calculations we omitted the year 1970 since the FES data for 1970 to which we had access are incomplete. As already mentioned in Chapter 2, the data for the first thirteen weeks of 1970 are missing.

Clearly, the aggregate labour supply function plotted in Figure 3 is a mixture of the labour supply functions of various groups of workers, i.e., the function is of the form $l(w) = \pi_1 l_1(w)+...+\pi_K l_K(w)$, where $l_i(\cdot)$ denotes the labour supply schedule of a suitably defined subgroup of workers; $\pi_i = \pi_i(w) \geq 0$ $(i=1,...,K)$ and $\pi_1+...+\pi_K=1$. Here, $\pi_i(w)$ represents the proportion of individuals of type i in the subgroup of workers receiving the wage rate w (presumably, the number K of groups is not small; see the remarks made at the beginning of Section 4 of Chapter 2). In principle, one could have used this relation and the curves given in Figure 6 in order to compute the aggregate labour supply function. We leave it to the reader to decompose the kernel regression estimator for a given sample $\{(y_i,w_i): i=1,...,n\}$ into a weighted sum of regression estimators for subsamples $\{(y_i,w_i): i \in I_j\}$ $(j=1, ...,K)$, where $I_j \cap J_k = \emptyset$, if $j \neq k$, and $I_1+...+I_K=\{1,...,n\}$.

Looking at Figure 6 and Table 1, we can say the following about the aggregate labour supply function: In the lower range of ρ essentially manual females are included in the average $l(w)$; around the maximum of $l(\cdot)$ at $w \approx$£3.5, we find many manual males, and in the upper range of ρ the function values $l(w)$ are determined by the labour supply of non-manual males.

We will now give a brief description of the labour supply functions drawn in Figures 6 and 7. The function values and sample percentiles given below are only approximate values. More precisely, the function values were computed from the diagrams in the two figures, and the sample percentiles were "guesstimated" using Table 1 and the density estimates given in Figure 7 of Chapter 2 (see pp. 103-107) [when this chapter was written, the author did not have access to the FES data]; w_x denotes the x-th percentile of the underlying gross wage rate distribution.

Table 2

Empirical Correlation between Labour Supply and Gross Wage Rate (1983)

	ALL WORKERS	FEMALES	MANUAL FEMALES	NON-MANUAL FEMALES	MALES	MANUAL MALES	NON-MAN. MALES	MANUALS	NON-MANUALS
COR	0.135	0.075	0.232	-0.035	-0.104	0.114	-0.177	0.379	0.074
Normal Weekly Hours in Excess of 30 (resp. ≤30)									
COR	-0.103	-0.183	-0.126	-0.156	-0.197	-0.080	-0.212	0.0004	-0.079
	(0.111)	(0.065)	(0.161)	(-0.001)	(0.228)	(0.363)	(0.095)	(0.210)	(0.047)
MEAN and STD of the Correlation Coefficients for 1971-85									
MEAN	0.153	0.052	0.161	-0.051	-0.089	0.064	-0.120	0.344	0.104
STD	0.024	0.037	0.065	0.044	0.024	0.033	0.043	0.031	0.032
Normal Weekly Hours in Excess of 30									
MEAN	-0.063	-0.148	-0.080	-0.121	-0.181	-0.098	-0.157	0.012	-0.014
STD	0.030	0.030	0.049	0.036	0.021	0.022	0.047	0.024	0.043

Table 3

Empirical Correlation between Net Earnings and Gross Wage Rate (1983)

	ALL WORKERS	FEMALES	MANUAL FEMALES	NON-MANUAL FEMALES	MALES	MANUAL MALES	NON-MAN. MALES	MANUALS	NON-MANUALS
COR	0.787	0.660	0.674	0.622	0.807	0.849	0.764	0.861	0.767
Normal Weekly Hours in Excess of 30 (resp. ≤30)									
COR	0.893	0.928	0.911	0.925	0.879	0.853	0.865	0.874	0.895
	(0.673)	(0.617)	(0.564)	(0.581)	(0.716)	(0.913)	(0.628)	(0.762)	(0.634)
MEAN and STD of the Correlation Coefficients for 1971-85									
MEAN	0.792	0.618	0.588	0.598	0.823	0.795	0.836	0.819	0.794
STD	0.045	0.067	0.071	0.079	0.050	0.046	0.056	0.039	0.057
Normal Weekly Hours in Excess of 30									
MEAN	0.897	0.925	0.879	0.928	0.884	0.822	0.895	0.852	0.918
STD	0.014	0.013	0.024	0.016	0.015	0.019	0.020	0.015	0.015

The labour supply function for **female workers** is decreasing on [0, $w_{0.05}$], increasing on [$w_{0.05}$,$w_{0.85}$] and decreasing on [$w_{0.85}$,$w_{0.99}$]. We have l(0)=27, l($w_{0.05}$)=22.5, l($w_{0.85}$)=33, and l($w_{0.99}$)=25. The labour supply function for **full-time female workers** is decreasing on [0,$w_{0.99}$] with l(0)=45, l($w_{0.05}$)=40.5, l($w_{0.50}$)=38.2, l($w_{0.90}$)=37.8 and l($w_{0.99}$)=36.8.

The labour supply function for **male workers** lies everywhere above that for female workers. The function is increasing on [0,$w_{0.30}$] and decreasing on [$w_{0.30}$,$w_{0.99}$]. The function values corresponding to 0, $w_{0.30}$ and $w_{0.99}$ are 36.6, 43.5 and 37.6, respectively. The labour supply function for **full-time male workers** lies above that for full-time female workers and is decreasing on [0,$w_{0.99}$]: l(0)=46, l($w_{0.05}$)=44.7, l($w_{0.90}$)=40.8 and l($w_{0.99}$)= 38.6.

The labour supply function for **manual female workers** is decreasing on [0,$w_{0.10}$] and increasing on [$w_{0.10}$,$w_{0.99}$]. We have l(0)=26.3, l($w_{0.10}$)= 20.6, l($w_{0.90}$)=30.6 and l($w_{0.99}$)=33. The labour supply function for **full-time manual female workers** is decreasing on [0,$w_{0.40}$], slightly increasing on [$w_{0.40}$,$w_{0.90}$] and almost flat on [$w_{0.90}$,$w_{0.99}$]; l(0)=44.3, l($w_{0.05}$)=42, l($w_{0.40}$)=39, l($w_{0.90}$)=39.7 and l($w_{0.99}$)=39.4.

The labour supply function for **manual male workers** lies entirely above that for manual female workers. The curve is increasing on [0,$w_{0.30}$] and slightly decreasing on [$w_{0.30}$,$w_{0.99}$]. The function values corresponding to 0, $w_{0.30}$ and $w_{0.99}$ are 31.2, 43.7 and 41.5, respectively (notice that the ordinary kernel estimate is rapidly decreasing on [0,0.7] (in £) with l(0)=40.) The labour supply curve for **full-time manual male workers** lies everywhere above that for full-time manual female workers and oscillates slightly around a gently decreasing function. We have l(0)=44.7, l($w_{0.05}$)= 44.2, l($w_{0.90}$)=43.3 and l($w_{0.99}$)=42.7.

The labour supply function for **non-manual female workers** lies below that for non-manual male workers and is decreasing on [0,$w_{0.01}$], increasing on [$w_{0.01}$,$w_{0.80}$] and decreasing on [$w_{0.80}$,$w_{0.99}$]; l(0)=27, l($w_{0.01}$)=24.5, l($w_{0.80}$)=33 and l($w_{0.99}$)=23.5. Since only 10 per cent of the manual females earn more than £2.6 per hour, the labour supply functions for the groups "non-manual female workers" and "all female workers" are approximately equal on [2.6,8] (in £); on [0,2.6] the function for the former group assumes larger values. The labour supply function for **full-time non-manual female workers** lies below that for full-time non-manual male workers and is decreasing on $[0,w_{0.99}]$. The curve has the same shape as that for the total population of full-time females; it lies, however, very slightly below the latter curve. The function values at 0, $w_{0.05}$, $w_{0.50}$, $w_{0.85}$ and $w_{0.99}$ are 45, 39.4, 37.5, 37.9, and 36.7, respectively.

The labour supply function for **non-manual male workers** is increasing on $[0,w_{0.15}]$ and decreasing on $[w_{0.15},w_{0.99}]$ with $l(0)=35.5$, $l(w_{0.15})=43$ and $l(w_{0.99})=36.5$. The labour supply function for **full-time non-manual male workers** has the same shape as that for "full-time male workers". The function values corresponding to 0, $w_{0.05}$, $w_{0.90}$ and $w_{0.99}$ are 48.2, 45.5, 39.7 and 38.5, respectively.

The labour supply function for **manual workers** lies almost entirely above that for "non-manual workers" and is decreasing on $[0,w_{0.05}]$, increasing on $[w_{0.05},w_{0.90}]$ and slightly decreasing on $[w_{0.90},w_{0.99}]$. We have $l(0)=28.2$, $l(w_{0.05})=25.6$, $l(w_{0.90})=43.1$ and $l(w_{0.99})=41.2$. The labour supply curve for **full-time manual workers** cannot be approximated by a decreasing function. The curve is decreasing on $[0,w_{0.10}]$, increasing on $[w_{0.10},w_{0.50}]$ and decreasing on $[w_{0.50},w_{0.99}]$; $l(0)=44.13$, $l(w_{0.10})=42.1$, $l(w_{0.50})=43.55$ and $l(w_{0.99})=42.75$.

The labour supply function for **non-manual workers** has the same shape as the aggregate labour supply function drawn in Figure 3. The curve is decreasing on $[0,w_{0.01}]$, increasing on $[w_{0.01},w_{0.60}]$ and decreasing on $[w_{0.60},w_{0.99}]$. The function values corresponding to 0, $w_{0.01}$, $w_{0.60}$ and $w_{0.99}$ are 28.9, 27.6, 38.2 and 34.5, respectively. The labour supply function for **full-time non-manual workers** has the same shape as the labour supply function for the total population of full-time workers plotted in Figure 4. The function is decreasing on $[0,w_{0.08}]$, gently increasing on $[w_{0.08},w_{0.50}]$ and decreasing on $[w_{0.50},w_{0.99}]$ with values $l(0)=45$, $l(w_{0.05})=40.6$, $l(w_{0.08})=39.5$, $l(w_{0.50})=40.6$ and $l(w_{0.99})=38.5$.

Turning to Figure 7, we see that the net earnings function for females lies entirely below that for males; and the net earnings function for manual (resp. non-manual) females lies everywhere below that for manual (resp. non-manual) males. Clearly, this is an obvious consequence of what has been said above about the labour supply curves. The estimated earnings function for women is downward sloping for gross wage rates near to the right endpoint of the w-axis of the corresponding diagram in Figure 7. However, when looking at Figure 7, it should not be forgotten that the range of the gross wage rate distribution for females is much smaller than that for males (Table 1). The sample contains only 30 (resp. 15) females (resp. full-time females) whose gross wage rate is greater than £8 (resp. £7), and none of the manual females earns more than £6.5 per hour.

The net earnings function for "manual workers" intersects that for "non-manual workers" twice, namely at $w \approx$ £1.9 and $w \approx$ £8.7. However, the relevant range of the gross wage rate distribution for "manual workers" ends at £6 (there are only 32 manual workers in the sample who earn more than £6 per hour). In the range between £1.9 and £6 the curve for "manual workers" lies above that for "non-manual workers". Clearly, this was to be expected since around 90 per cent of the full-time manual males are located in this interval.

The correlation coefficients of the data from which Figure 7 was constructed are given in Table 3 above (see also the tables in the Appendix). Observe that excluding part-time workers from a sample leads to a larger correlation coefficient; the correlation coefficients for part-time workers are substantially lower than those for full-time workers (the only exception is the group "manual male workers"). As the second part of Table 3 shows, the correlation coefficients in the samples of full-time workers are remarkably stable across the years.

Notice that in Table 3 the exclusion of part-time workers from the annual samples leads to a significant reduction in the spread of the correlation coefficients around its arithmetic mean. This is not the case in Table 2. Looking along the third row from the bottom of Table 2 (resp. Table 3), we see that the standard deviations are larger in the case of net earnings. Comparing the last row of Table 2 with the last row of Table 3, it is interesting to observe that the standard deviations are now larger in the case of labour supply.

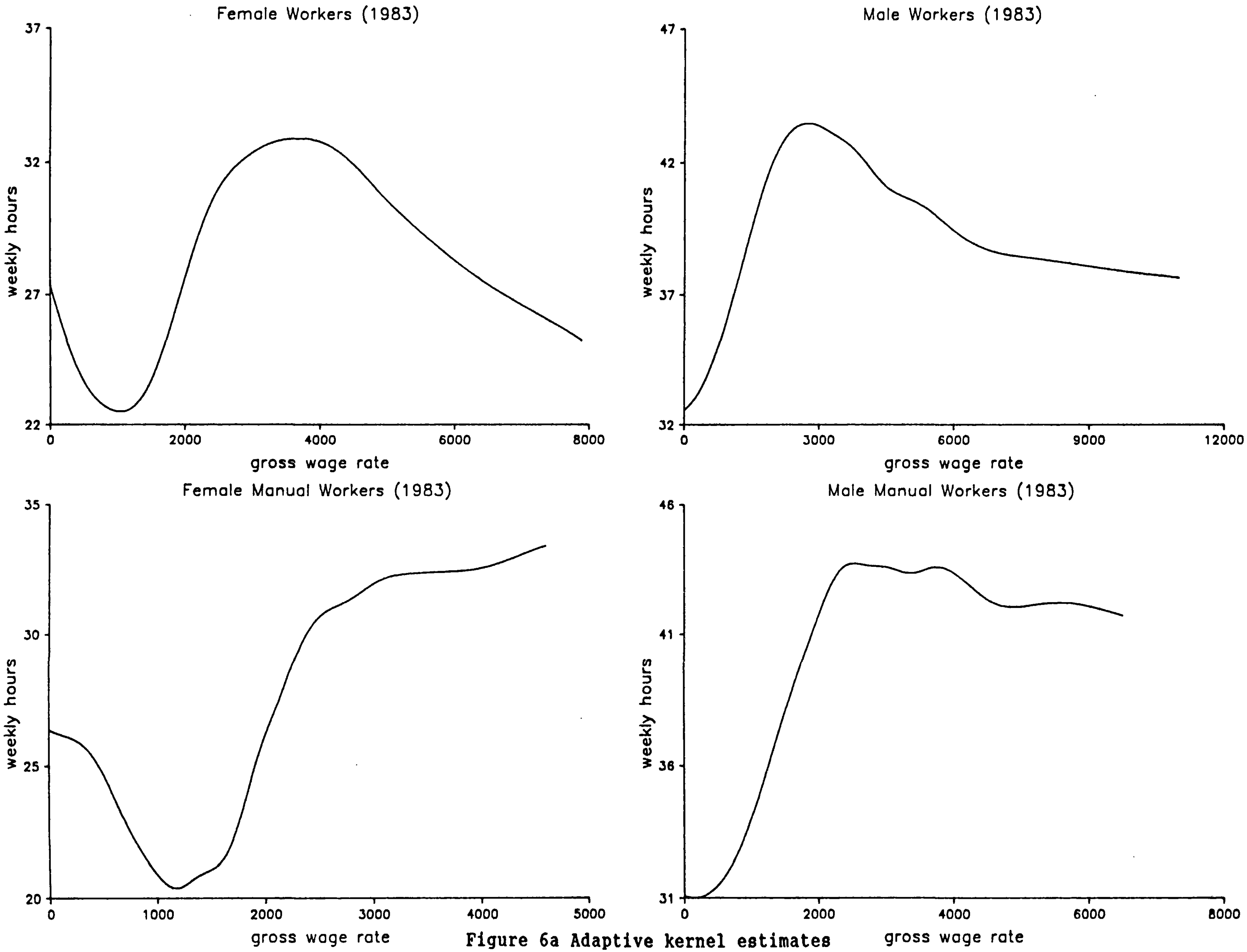


Figure 6a Adaptive kernel estimates

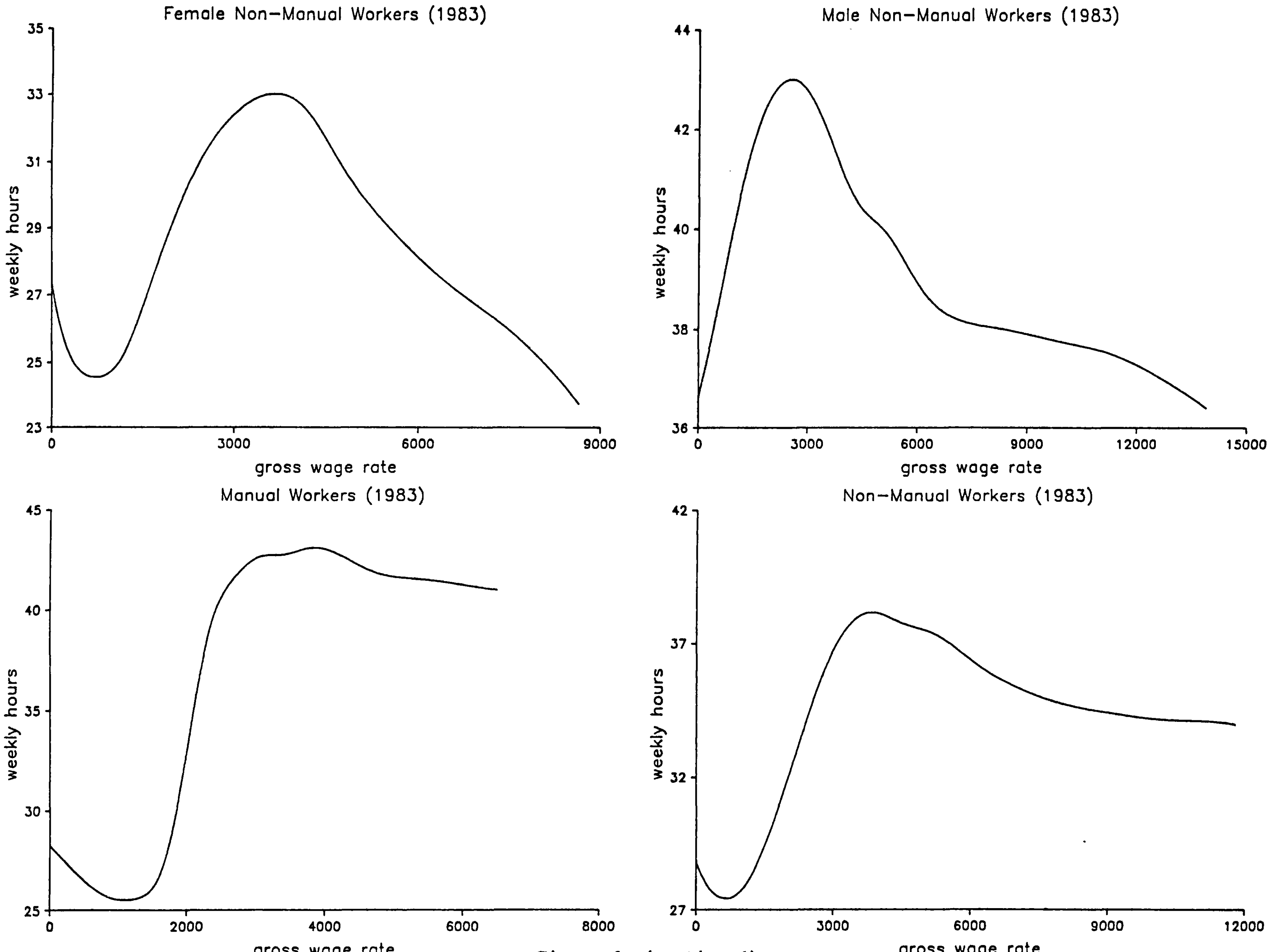


Figure 6a (continued)

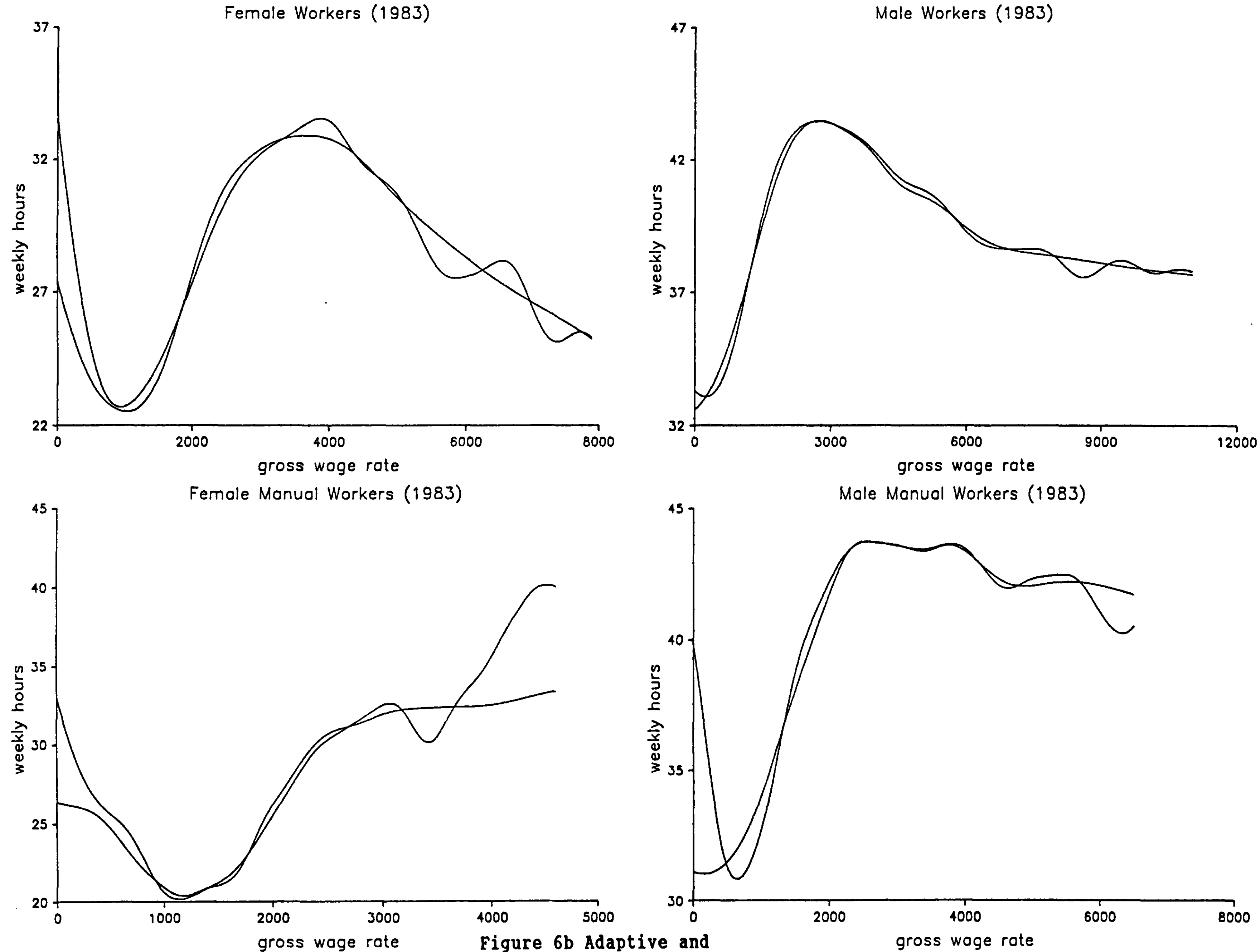


Figure 6b Adaptive and ordinary kernel estimates

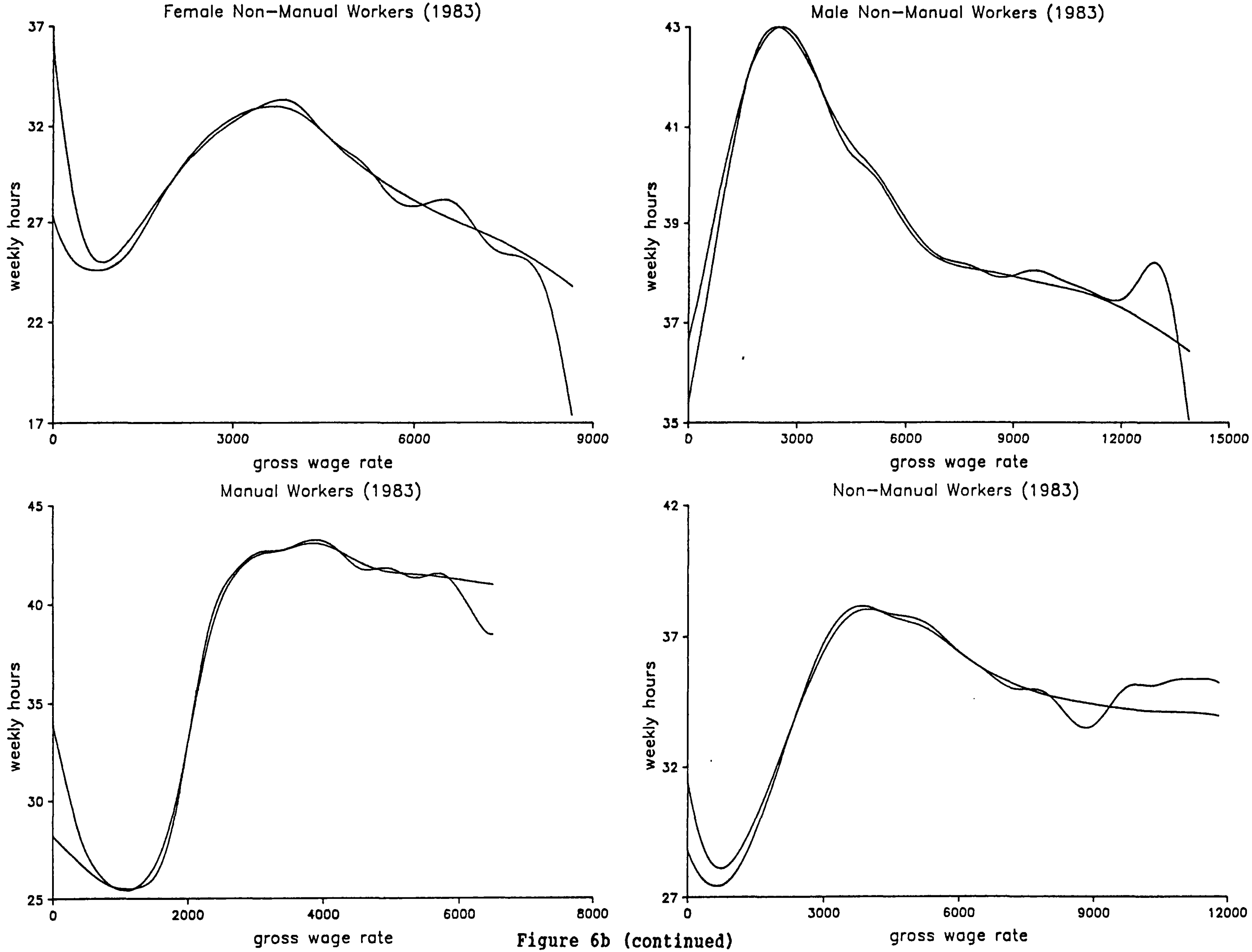


Figure 6b (continued)

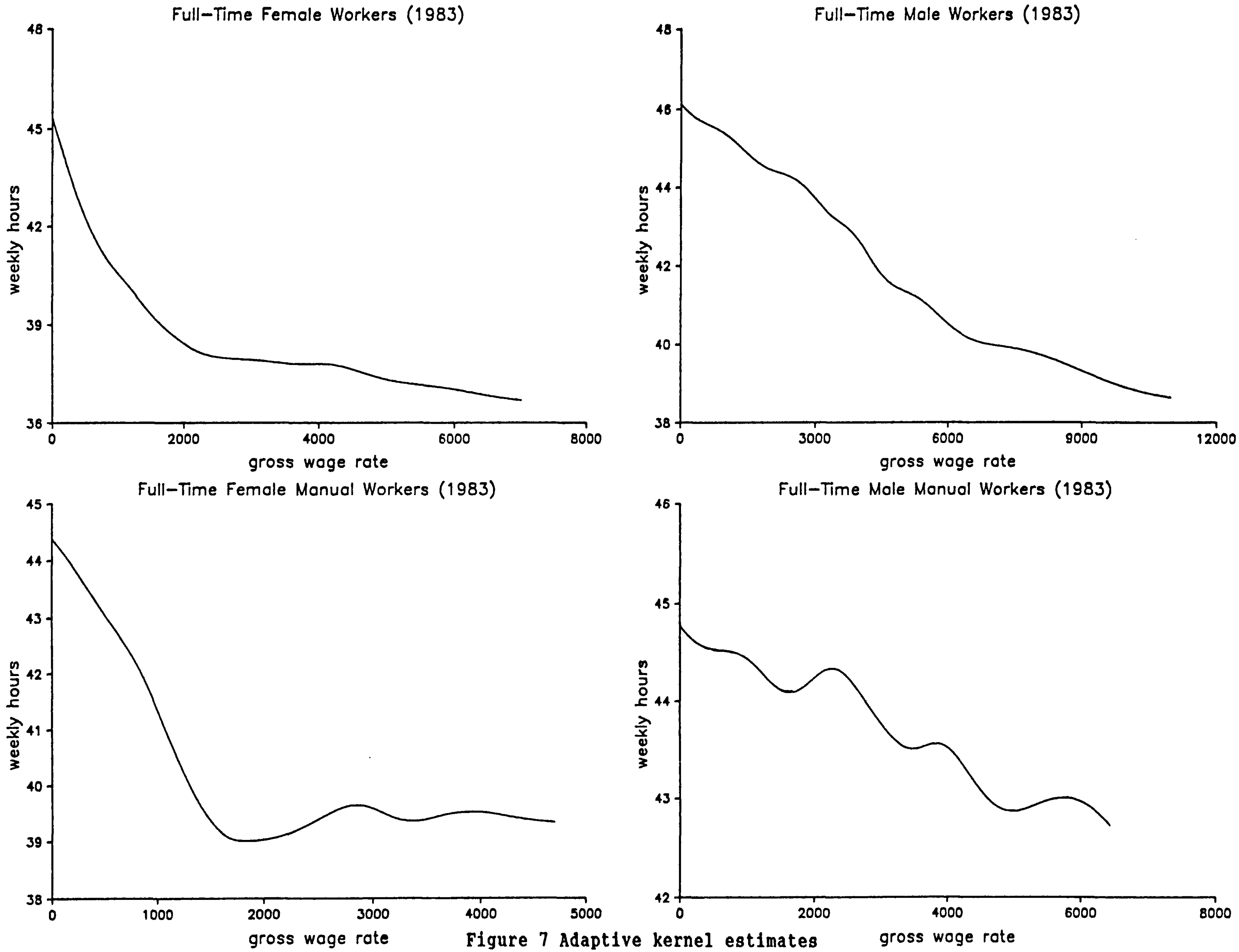


Figure 7 Adaptive kernel estimates

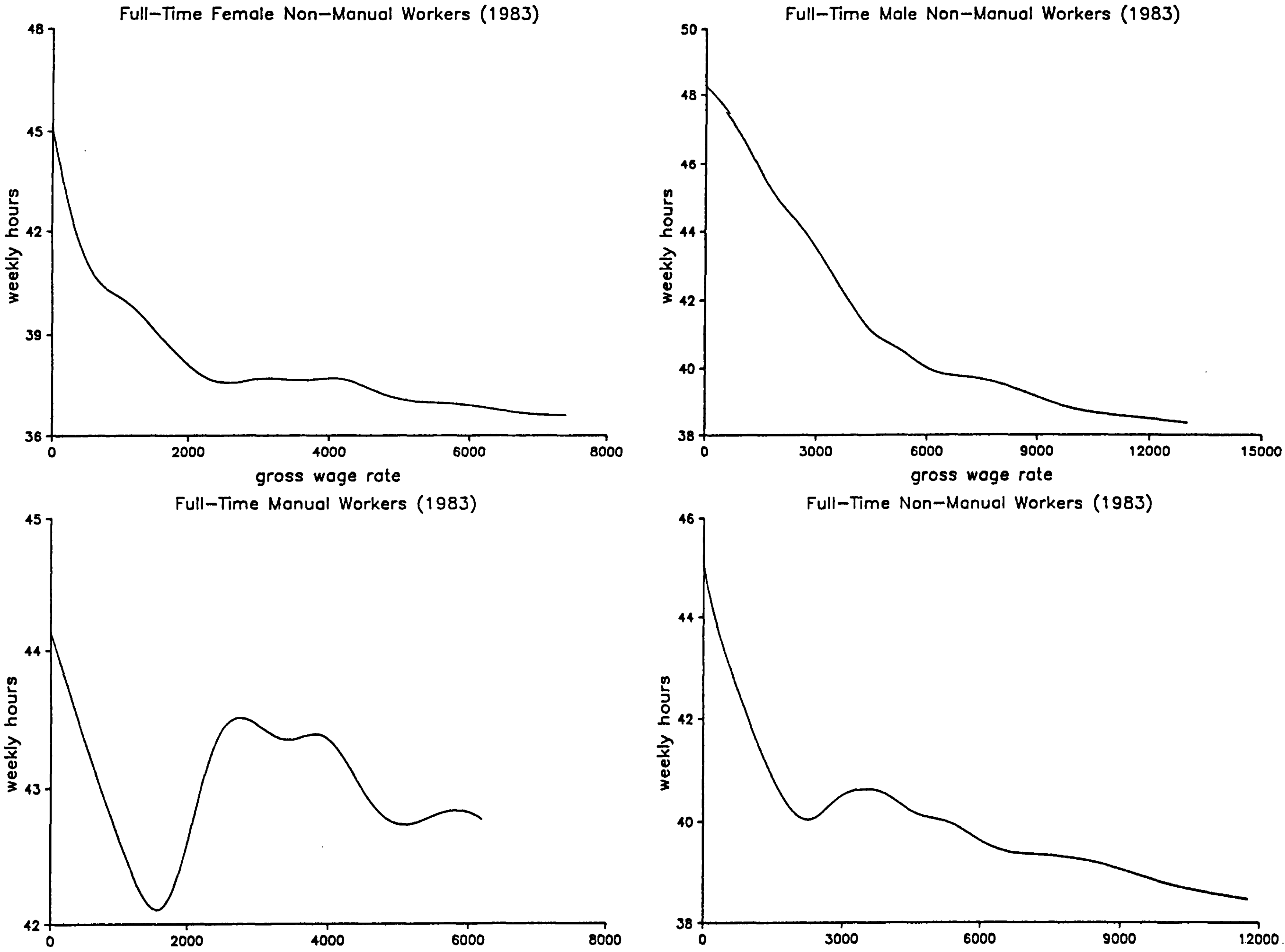
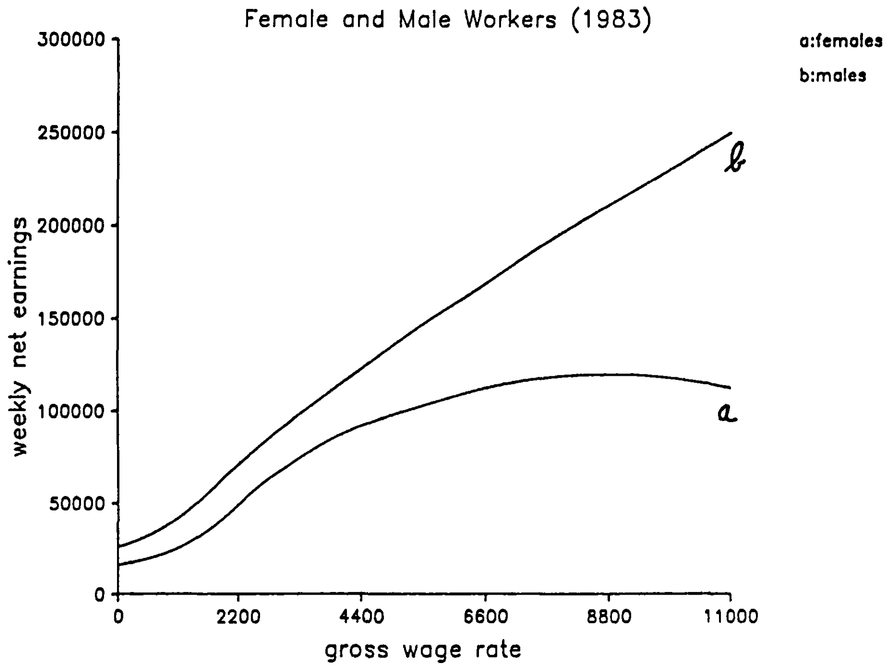
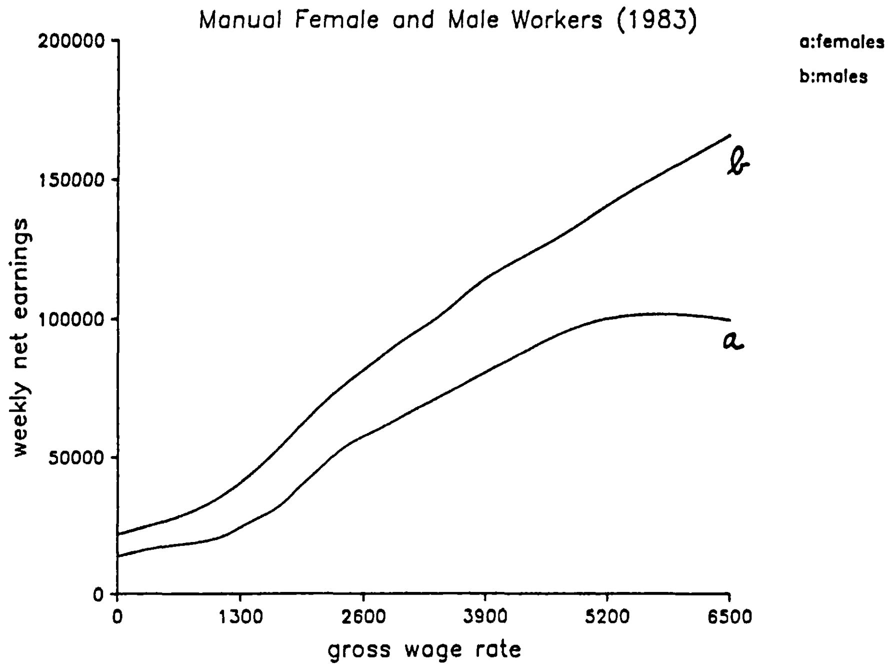
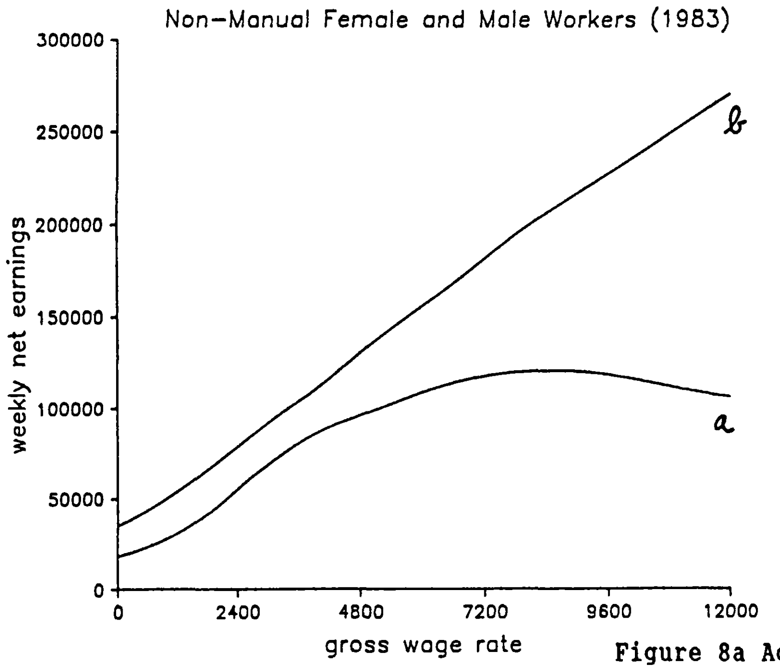


Figure 7 (continued)







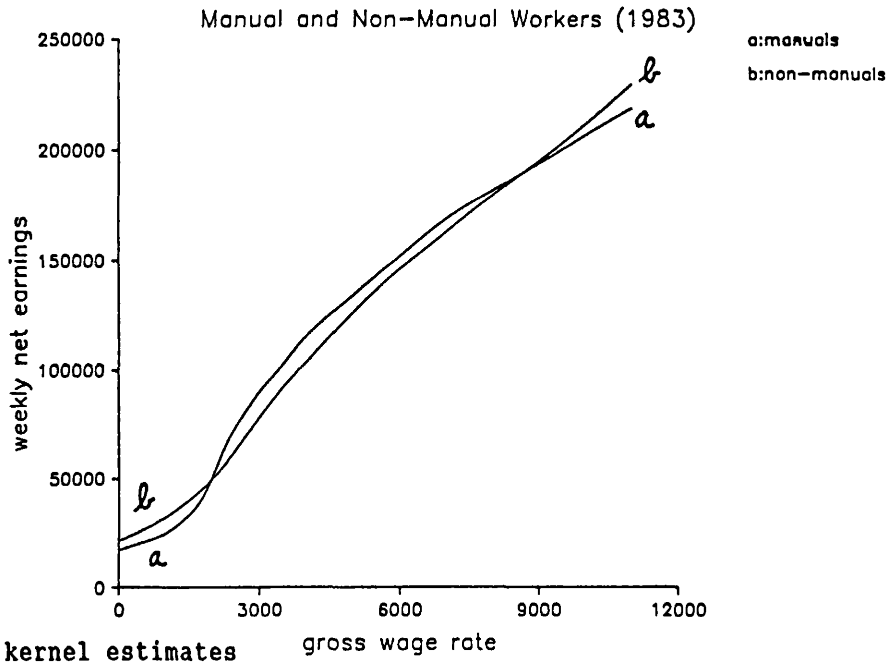


Figure 8a Adaptive kernel estimates

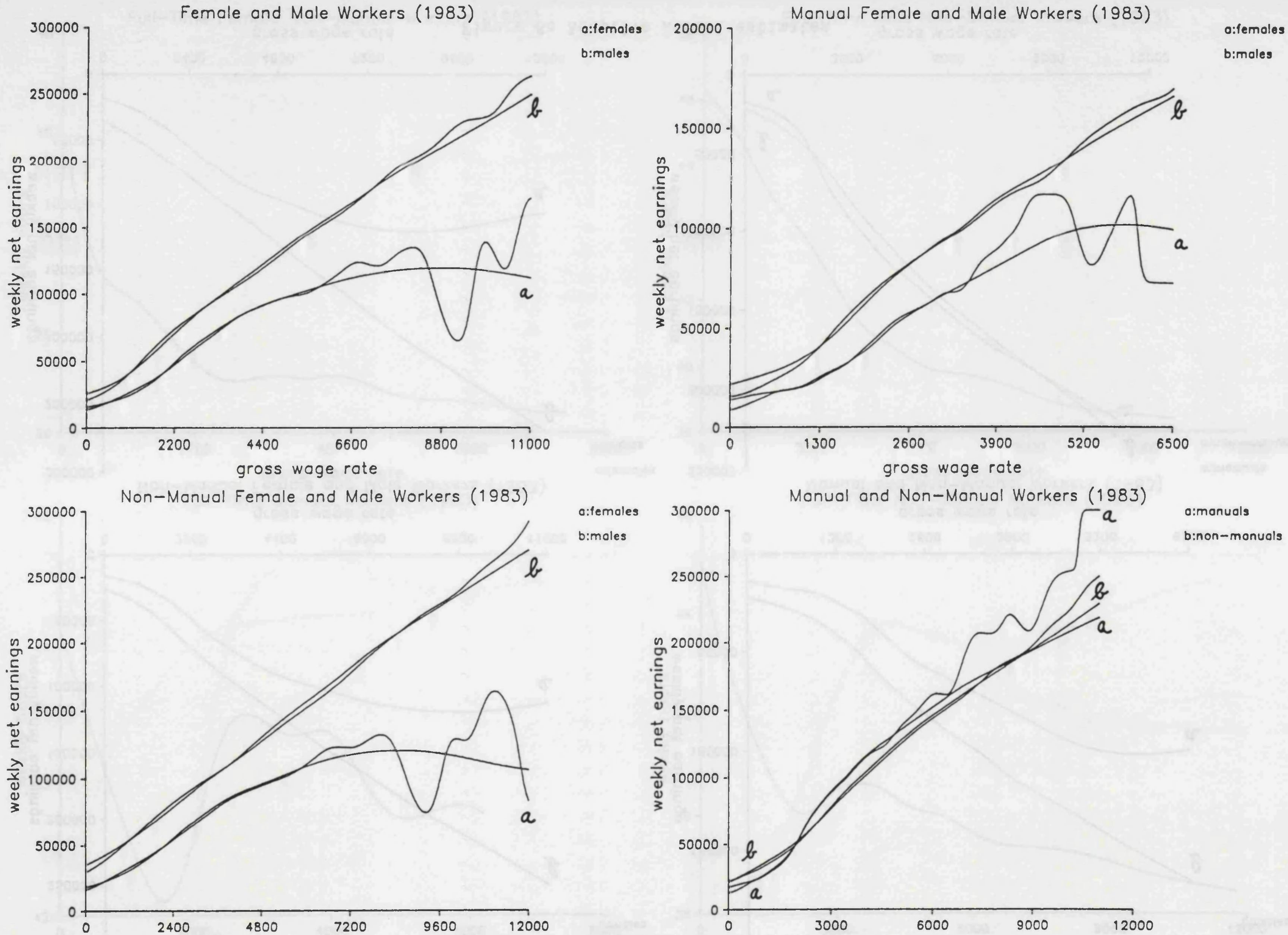


Figure 8b Adaptive and ordinary kernel estimates

5. Labour Supply and Net Earnings Functions over Time

We now want to explore whether or not the aggregate labour supply (resp. net earnings) function did change in the period from 1971 to 1985. The section is related to Härdle and Jerison (1988) who estimate Engel curves; the authors remark in their introduction: "the theoretical and applied demand literature has devoted little attention to the evolution of cross section Engel curves". To our knowledge this applies also to labour supply functions. We begin with some general remarks.

Recall that the two components of our simple labour supply model are the statistical labour supply function $l(\cdot)$ and the density ρ of the distribution of gross wage rates. One can therefore decompose a change in the gross wage rate elasticity into a change in $l(\cdot)$ and ρ. Let $l_t(\cdot)$ and ρ_t denote, respectively, the labour supply function and the gross wage rate density of a particular population of workers in period t. Per capita labour supply of this population at time t is then given by $L_t=\int l_t(w)\rho_t(w)dw$, and the rate of change of L_t resulting from a small absolute increase in w is given by $L'_t=\int l'_t(w)\rho_t(w)dw$. If $l_t(\cdot)$ and ρ_t change over time, then L_t and L'_t will in general also change. The change in L'_t between, say, t=0 and t=1 can be written in the form

$$L'_1 - L'_0 = \int[l'_1(w) - l'_0(w)]\cdot\rho_1(w)dw + \int l'_0(w)[\rho_1(w) - \rho_0(w)]dw.$$

The first term after the equality sign represents the effect of a change in labour supply behaviour on L'_t, and the second term represents the effect of a change in the wage rate distribution. If $L'_1-L'_0=0$, then we are faced with two completely offsetting effects on L'_t. If one is only interested in the labour supply behaviour of workers in a particular wage rate class, say, [a,b], one has to replace ρ_t by the on [a,b] truncated density

$$\tilde{\rho}_t(w) = \begin{cases} \dfrac{\rho_t(w)}{\int_a^b \rho_t(w)dw}, & \text{if } w \in [a,b] \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding expression for the change in the gross wage rate elasticity between t=0 and t=1 is somewhat more complicated. It is, however, a straightforward exercise to check that this change can be written as

$$\frac{1}{L_1}\cdot A + \frac{1}{L_0}\cdot B + \frac{1}{L_0 L_1}\cdot C\cdot(D_1 + D_2),$$

where

$$A = \int[l_1'(w) - l_0'(w)]\cdot w\cdot\rho_1(w)dw,$$

$$B = \int l_0'(w)\cdot w\cdot[\rho_1(w) - \rho_0(w)]dw,$$

$$C = \int l_0'(w)\cdot w\cdot\rho_1(w)dw,$$

$$D_1 = \int[l_0(w) - l_1(w)]\cdot\rho_0(w)dw$$

and

$$D_2 = \int l_1(w)\cdot[\rho_0(w) - \rho_1(w)]dw.$$

If $l_t(w)=l(w)$ for all t, then A and D_1 vanish and one could use cross-section labour supply curves in order to compute the effect of changes in the wage rate distribution on hours of work. Of course, the labour supply curve in period t depends on various exogenous parameters. In particular, it depends upon the consumer goods prices and the tax function in this period. Supposing the curves shift over time, then an important question for the economist is whether the shifts can be plausibly explained by observed changes in the relative prices. If the curves shift substantially even during periods of stable prices and wage rates, then predictions based on cross-section labour supply curves will be very unreliable.

Notice that the set of individuals constituting the population will typically change over time. However, our model is set up in such a manner that it does not refer to individual members of the population. The relevant question within the present analysis is whether the demographic composition of the population changed across the years. In Chapter 2 we have seen that the composition of the labour force changed considerably during the 1970s. We will return to this point later.

Recall that the mean and the variance of the earnings data increase steadily from 1970 to 1985 (see Table 1 in the Appendix). In order to explore the evolution of the labour supply functions, we therefore have to normalise the explanatory variable. As in the case of density estimation, we divided all gross wage rates in the annual samples by its arithmetic mean. Thus, one obtains the "normalised" labour supply curve for the year 1983 by simply rescaling the w-axis of Figure 3. Clearly, what has been said above applies also to mean normalised labour supply functions.

Let μ_t be the mean of the gross wage rate distribution in period t. Suppose the mean labour supply of workers receiving the wage rate w depends only on their wage rate relative to the mean, i.e., for all (w,t) and (w',t') we have

$$\text{(I)} \quad \frac{w}{\mu_t} = \frac{w'}{\mu_{t'}} \quad \text{implies } l_t(w) = l_{t'}(w').$$

Then the functions $l_t(\cdot)$ admit the representation $l_t(w)=l(w/\mu_t)$. Since any labour supply function $l_t(\cdot)$ can be written as $l_t(w)=\tilde{l}_t(w/\mu_t)$, one tests (I) by regressing labour supply on the normalised wage rate w/μ_t and checking whether one obtains essentially the same curve in each year. Notice, if (I) holds, then the elasticity of labour supply is determined by the distribution of normalised gross wage rates, i.e., by the density function

$$\tilde{\rho}_t(w) = \mu_t \cdot \rho_t(\mu_t \cdot w).$$

In the case of net earnings we also have to normalise the dependent variable. The (mean) normalised net earnings function for year t is obtained from the nominal curve, $b_t(\cdot)$, by dividing both the dependent and the independent variable by its mean value. Let $\rho_t(b,w)$ denote the density of the joint distribution of net weekly earnings and gross wage rates at time t, and let B_t denote the mean of the marginal distribution of net earnings. We say that the net earnings curves exhibit *mean normalised invariance* with respect to a change in the density function $\rho_t(b,w)$ if there exists a function $b(\cdot)$ such that

$$\frac{1}{B_t} \cdot b_t(w) = b\left(\frac{w}{\mu_t}\right) \quad \text{for all } t.$$

Let us now turn to the estimations. We estimated the labour supply (resp. net earnings) function on the entire sample of workers for each odd numbered year from 1971 until 1985. The results are given in Figures 9 and 10 on pages 206-211. Figures 9a and 9b show adaptive kernel estimates of the aggregate labour supply functions; in Figure 9c each adaptive kernel estimate is plotted together with the corresponding ordinary kernel estimate. As always, the ordinary kernel method generates curves which are unstable in the upper tail of the underlying gross wage rate distribution, i.e., in the range of normalised wage rates which are greater than 1.6. For the given value of the smoothing parameter (h=0.12), the method produces, however, reasonable results in the lower range of ρ_t. Observe that in Figure 9c the deviations (of the ordinary from the adaptive kernel estimates) are smaller than in Figure 6b. (We could improve in Figure 6b the performance of the ordinary kernel method by selecting somewhat larger values for the smoothing parameter.)

Figure 10 presents adaptive kernel estimates of the normalised net earnings functions for the eight years. As in Figure 9, the curves were computed with h=0.12. The OLS regression line for 1983 plotted in Figure 10 is given by (the figures in brackets are the standard errors):

$$b(w) = \underset{(0.009)}{0.194} + \underset{(0.008)}{0.806}\cdot w.$$

Sample percentiles for the normalised gross wage rates are given in Table 2 of Chapter 2 (page 125). The density estimates corresponding to Figures 9 and 10 are plotted in Figure 12 of Chapter 2 (pages 136-139). Observe that we used for the density functions a smaller window width, namely h=0.08. The choice h=0.12 would have obscured the fine structure of the data in the main body of the distributions. From the point of view of the labour supply (resp. net earnings) elasticity, however, it does not matter very much whether we compute the normalised gross wage rate densities with h=0.08 or h=0.12.

In the following we will use the term "wage rate" as an abbreviation for "normalised gross wage rate". The reader may find it useful to look first through pages 203-205 which provide a discussion of the estimates.

As we see, the shape of the regression functions is remarkably stable across the years. The labour supply functions are over the interval [$w_{0.05}$, $w_{0.99}$] first strongly increasing and then decreasing. Since around 45 per cent of the wage rates are contained in the interval [0.4,0.9], 25 per cent in the interval [0.9,1.2] and 20 per cent in the interval [1.2,2.0] (Table 2, p. 125), we can infer from Figure 9 that a small rise in all wage rates would have implied an increase in aggregate labour supply in each year (if the model is correct and if the data are representative). In other words, when looking at the total population of workers, it appears that the labour market fulfils the "law of supply".

The net earnings functions are strictly increasing on [0,3.2], i.e., the decrease in labour supply in the upper range of the wage rate distribution does not lead to a decrease in disposable labour income. The curves are convex over the interval [0,1.1] and concave over [1.1,1.6].

In fact, there is little to say about the net earnings functions. Figure 10 shows that the data support impressively well the hypothesis of mean normalised invariance during the years 1971-85. In the central part of the wage rate distributions, i.e., in the interval [0.4,1.6], the curves are almost indistinguishable by eye. The picture does not change for wage rates near zero. In the upper range of ρ_t, however, where relatively few observations are distributed over a large interval, the curves spread out.

The labour supply functions are less stable during the years from 1971 to 1985. As can be seen in Figure 9a, the curves shift considerably in the first half of this period. The labour supply curves for the years 1981, 1983 and 1985 are almost equal over the interval [0.5,2.0], i.e., in the range from around the 10th to the 95th percentile of the underlying wage rate distributions. In the range from around the 75th to the 95th percentile they differ, however, substantially from the curve for 1979. Indeed, the latter curve would fit much better into the first diagram of Figure 9a (see also Figure 9b).

It is interesting to observe that during the 1970s the function values decreased while the shape of the labour supply functions did not change over the sample period. Supposing the labour supply function at time t can be written as $l_t(w) = l(w) + a_t$ with $a_t < a_{t-1}$ for all $t \leq t^*$, then the elasticity of labour supply will be, of course, increasing until t*. Hence, Figure 9 suggests that the elasticity increased during the 1970s. However, in Section 6 we will see that the difference between the smallest elasticity (in 1970) and the largest (in 1977) is only around 0.06.

Recall from Chapter 2 that during the 1970s the composition of the labour force changed considerably. More precisely, the composition of the annual FES samples suggests the following (see the tables on pages 127-9): Firstly, the proportion of women in the labour force grew from around 0.385 (in 1971) to 0.426 (in 1979). Secondly, there was a general switch from manual to non-manual occupations leading to an increase in the proportion of non-manual workers from around 0.409 (in 1971) to 0.462 (in 1979).

Since females work less hours per week than males and non-manuals less hours than manuals (see also the tables in the Appendix), it is reasonable to conclude that the observed shifts in the labour supply curves are attributable to these two changes in the composition of the annual samples. Of course, the explanation is unsatisfactory. It would be interesting to have a closer look at the determinants of the changes. In particular, the dependence of labour supply on the consumer goods prices and the tax function should be explored.

We remark that the FES data exhibit significant changes in the relative prices over the sample period (see Härdle and Jerison, 1988). Furthermore, in 1973 a new tax system was introduced with a higher exemption level which may have encouraged part-time work (the proportion of part-time females in the annual samples increases between 1971 and 1985 by 4.6 percentage points). The national insurance system changed also in the period. From 1975 onwards only those earning more than a specified amount had to pay contributions. Finally, the position of women in the labour market was changed by the Equal Pay Act of 1970, the Sex Discrimination Act of 1975 and the Employment Protection Acts of 1975, 1978, 1980 and 1982. As already mentioned in Chapter 2, a recent study on the rise in female employment during the 1970s is Gomulka and Stern (1989); see also note 9 in Chapter 2.

Discussion of the Estimated Curves

Looking at the first diagram of Figure 9a, we see that the labour supply curves for 1971 and 1973 lie entirely above that for 1977. Over the interval [0.4,3.2] the latter curve lies also almost entirely below that for 1975 (the labour supply function for 1977 assumes only on a small interval around w=0.93 slightly larger values). For wage rates smaller than 0.4 we have

$$l_{1975}(w) < l_{1977}(w) < l_{1971}(w) < l_{1973}(w).$$

In the range of wage rates between 0.5 and 2.0 the labour supply curve for 1971 lies almost everywhere above those for 1973 and 1975. The function for 1973 assumes only on a small interval around w=1.05 slightly larger values, and as long as the wage rate is not larger than 1.8, the function for 1975 does not assume larger values than that for 1973; over the interval [1.2,1.8] the graphs of the latter two functions are indistinguishable by eye. Hence, on [0.5,1.8]

$$l_{1977}(w) < l_{1975}(w) \leq l_{1973}(w) < l_{1971}(w).$$

For instance, at w=0.8 we have $l_{1971} \approx l_{1977}+2.75$. For wage rates greater than 2.0 we have $l_{1975}(w) > l_{1971}(w)$. In the range between 2.0 and 2.8 the labour supply curve for 1971 still lies above that for 1973. The picture changes as the wage rate approaches 3.2. The labour supply function for 1973 is increasing on [2.7,3.2], intersects that for 1971 at w≈2.8 and assumes at w=3.2 a larger value than that for 1975.

Looking at the period 1979-85, we see that there is no labour supply curve that lies entirely above or below another curve. Over the interval [0,0.5] the labour supply functions for 1981, 1983 and 1985 are first decreasing and then increasing; the former two curves differ only very slightly and lie below that for 1985. The function for 1979 is strictly increasing on [0,0.5] with $l_{1979}(0) < l_{1981}(0)$, $l_{1979}(0.4) > l_{1985}(0.4)$ and $l_{1979}(0.5) \approx l_{1985}(0.5)$. Notice that only the functions for 1981, 1983, 1985 and 1973 are decreasing on [0,0.3]; in the interval [0,0.5] the curve for 1973 lies above those for 1979-85.

Contrary to the period 1971-77, the curves in the second diagram of Figure 9a differ only very slightly in the range of wage rates between 0.5 and 0.9. Over the interval [0.5,0.9] we have

$$l_{1983}(w) < l_{1981}(w) < l_{1985}(w).$$

The curve for 1979 intersects the other three curves over [0.5,0.6] and over [0.82,0.90], approximately, so that it lies between 0.6 and 0.82 somewhat below that for 1983. Over the interval [1.0,2.2] the curve for 1979 lies considerably above those for 1981, 1983 and 1985.

As long as the wage rate is not larger than 1.9 the latter three curves differ only very slightly, but from then onwards they deviate more and more from each other. In the range between 1.2 and 1.7 we have

$$l_{1985}(w) < l_{1981}(w) < l_{1983}(w).$$

Over the interval [1.9,3.2] mean labour supply is smallest in 1983. For wage rates greater than 2.2 we have

$$l_{1983}(w) < l_{1981}(w) < l_{1985}(w),$$

and the difference between $l_{1983}(w)$ and $l_{1985}(w)$ increases as the wage rate approaches 3.2. The labour supply function for 1985 is increasing on [2.3,3.2] and that for 1981 is increasing on [2.5,3.2], while the functions for 1979 and 1983 are decreasing on the interval [1.2,3.2]. It is interesting to observe that the graphs of the latter two functions are almost parallel in this range with

$$l_{1979}(w) \approx l_{1983}(w) + 1.3.$$

The labour supply function for 1985 (resp. 1981) intersects that for 1979 at $w \approx 2.4$ (resp. $w \approx 2.65$), so that we have for wage rates greater than 2.7

$$l_{1983}(w) < l_{1979}(w) < l_{1981}(w) < l_{1985}(w).$$

Looking at the first diagram of Figure 10a, we see that the net earnings curves for 1971, 1973, 1975 and 1977 start to spread out for wage rates greater than 2.6. For wage rate greater than 2.8 we have

$$b_{1973}(w) > b_{1971}(w) > b_{1975}(w) > b_{1977}(w).$$

On [0,0.5] we have $b_{1971}(w) < b_{1977}(w)$; in the range between 0.1 and 0.5 the curves for 1973 and 1975 are approximately equal and lie very slightly below that for 1977. However, as $w \to 0$, $b_{1975}(w)$ approaches $b_{1971}(0)$, and $b_{1973}(w)$ approaches $b_{1977}(0)$. In the interval [0.55,0.65] the net earnings curve for 1977 intersects those for 1971, 1973 and 1975; over [0.65,0.85] we have

$$b_{1977}(w) \approx b_{1975}(w) < b_{1973}(w) \approx b_{1971}(w),$$

and over [0.85,1.05]

$$b_{1975}(w) < b_{1977}(w) \approx b_{1971}(w) < b_{1973}(w).$$

In the range of wage rates between 1.05 and 1.6 the four curves are indistinguishable by eye. For wage rates between 1.6 and 1.9 we have

$$b_{1977}(w) \approx b_{1975}(w) \approx b_{1973}(w) < b_{1971}(w).$$

Over the interval [1.9,2.4] the curves for 1973 and 1977 are still approximately equal and lie below that for 1971. However, the net earnings curve for 1975 intersects at $w \approx 2.05$ that for 1971 and lies in the interval [2.1,2.4] above the latter curve. The last point of intersection is at $w \approx 2.5$.

As was to be expected from the behaviour of the labour supply curves, the net earnings curves for 1979, 1981, 1983 and 1985 begin to fan out already at $w \approx 1.9$. For wage rates greater than 1.9 the 1979 curve lies below the other three curves. In the range between 1.6 and 3.2 the net earnings functions for 1981 and 1983 do not differ very much but that for 1981 assumes in general somewhat larger values. On [1.95,2.4] we have

$$b_{1985}(w) < b_{1983}(w) < b_{1981}(w).$$

At $w \approx 2.4$ the net earnings function for 1985 intersects those for 1981 and 1983 and assumes for wage rates greater than 2.4 considerably larger values than the other three functions. Notice that the curve lies in the interval [2.6,3.2] even above the OLS regression line for 1983. In the range between 0.1 and 0.45 we have

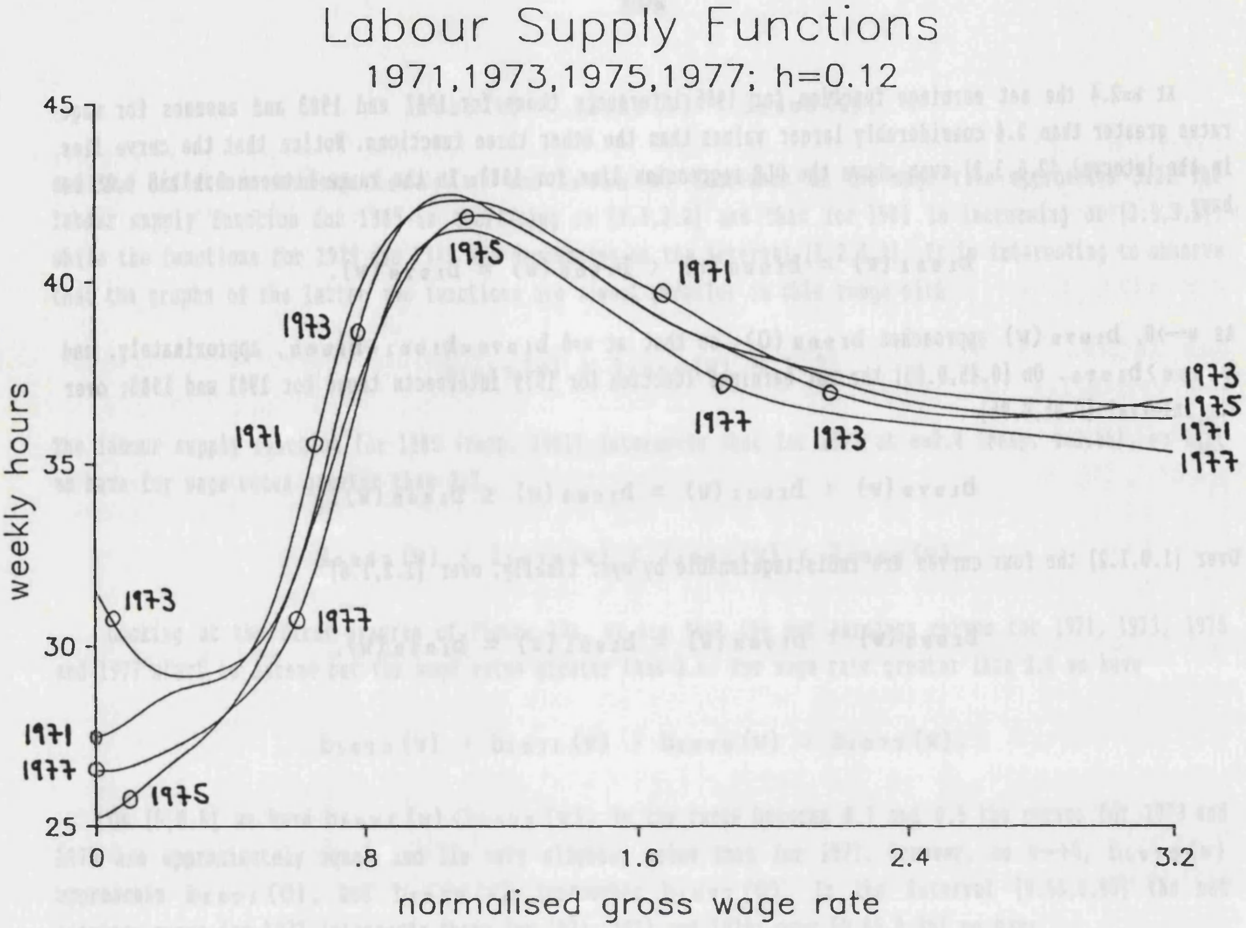
$$b_{1981}(w) \approx b_{1983}(w) < b_{1985}(w) \approx b_{1979}(w).$$

As $w \to 0$, $b_{1979}(w)$ approaches $b_{1983}(0)$, so that at $w=0$ $b_{1979}=b_{1981}=b_{1983}$, approximately, and $b_{1985}>b_{1979}$. On [0.45,0.55] the net earnings function for 1979 intersects those for 1981 and 1983; over the interval [0.55,0.95]

$$b_{1979}(w) < b_{1981}(w) \approx b_{1983}(w) \leq b_{1985}(w).$$

Over [1.0,1.2] the four curves are indistinguishable by eye; finally, over [1.2,1.6]

$$b_{1985}(w) < b_{1983}(w) \approx b_{1981}(w) \approx b_{1979}(w).$$



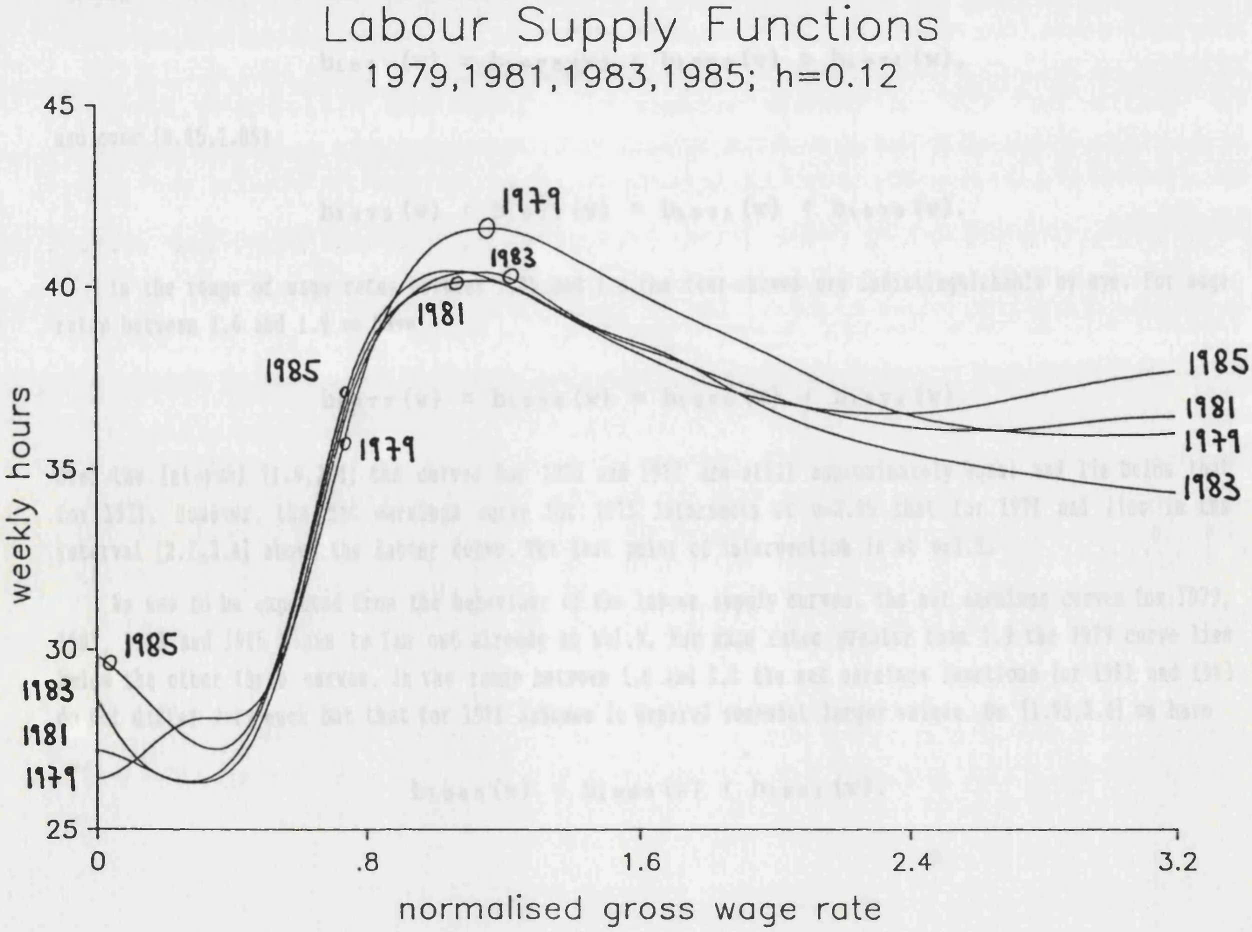


Figure 9a Adaptive kernel estimates

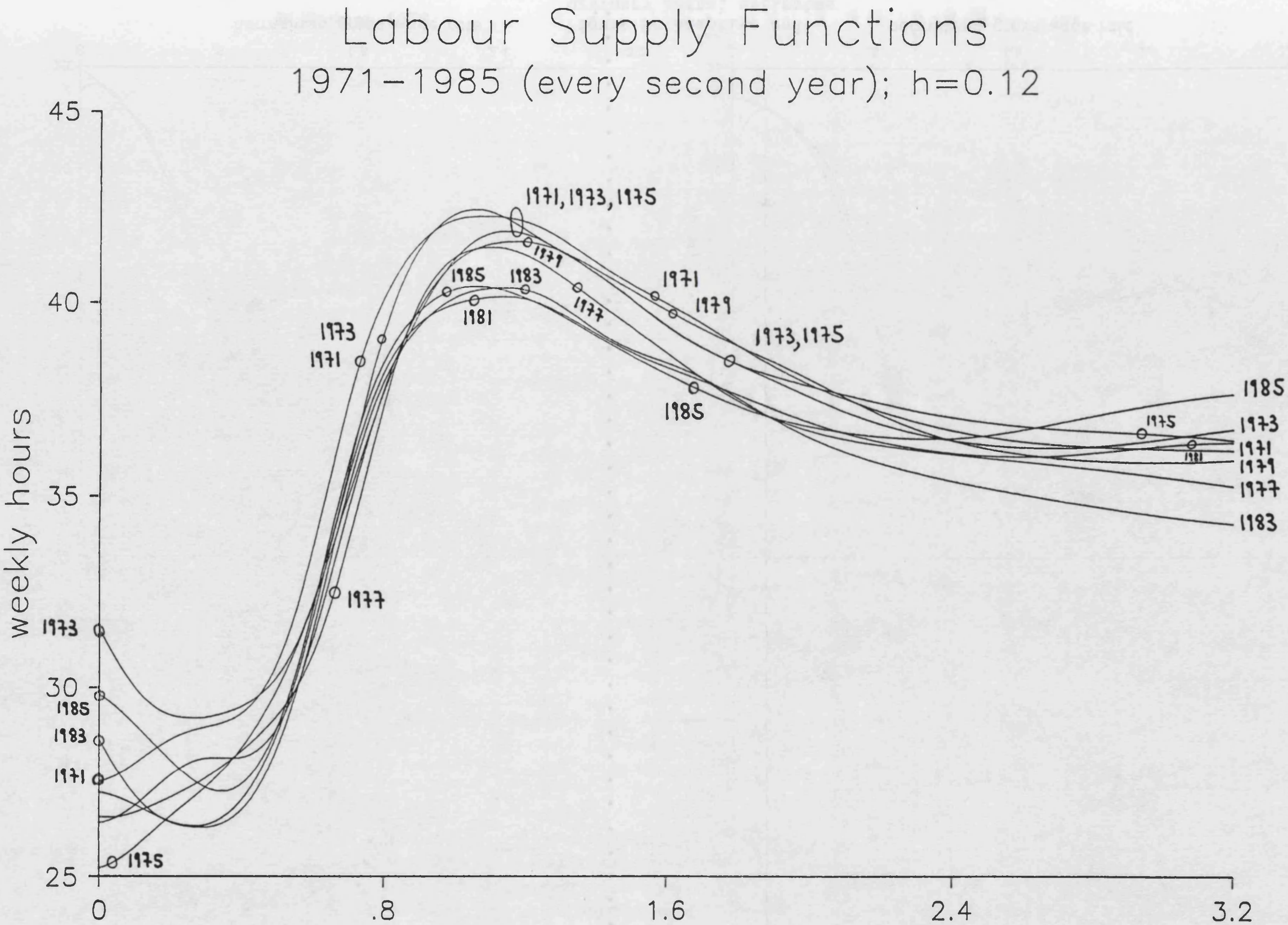


Figure 9b
Adaptive kernel estimates

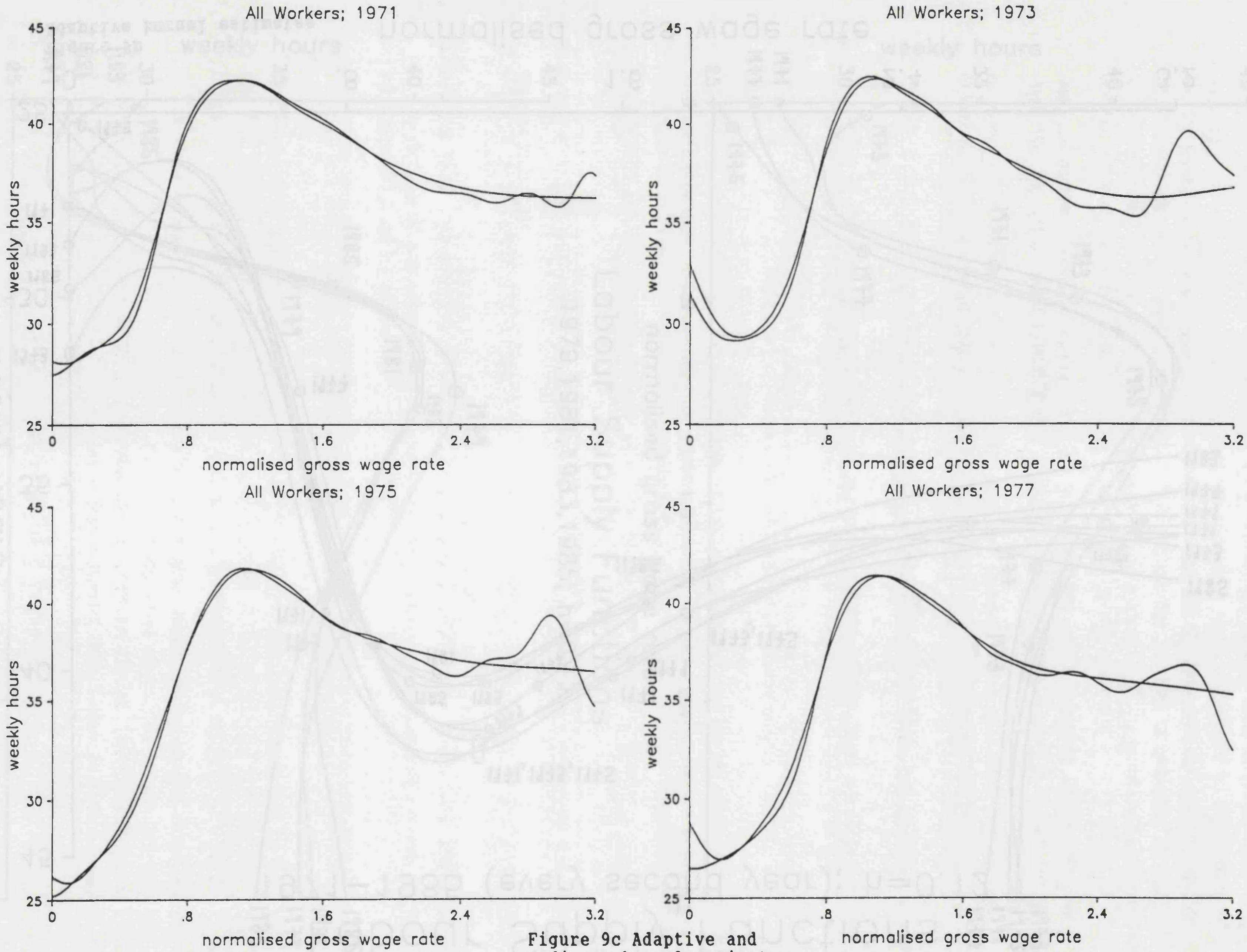


Figure 9c Adaptive and ordinary kernel estimates

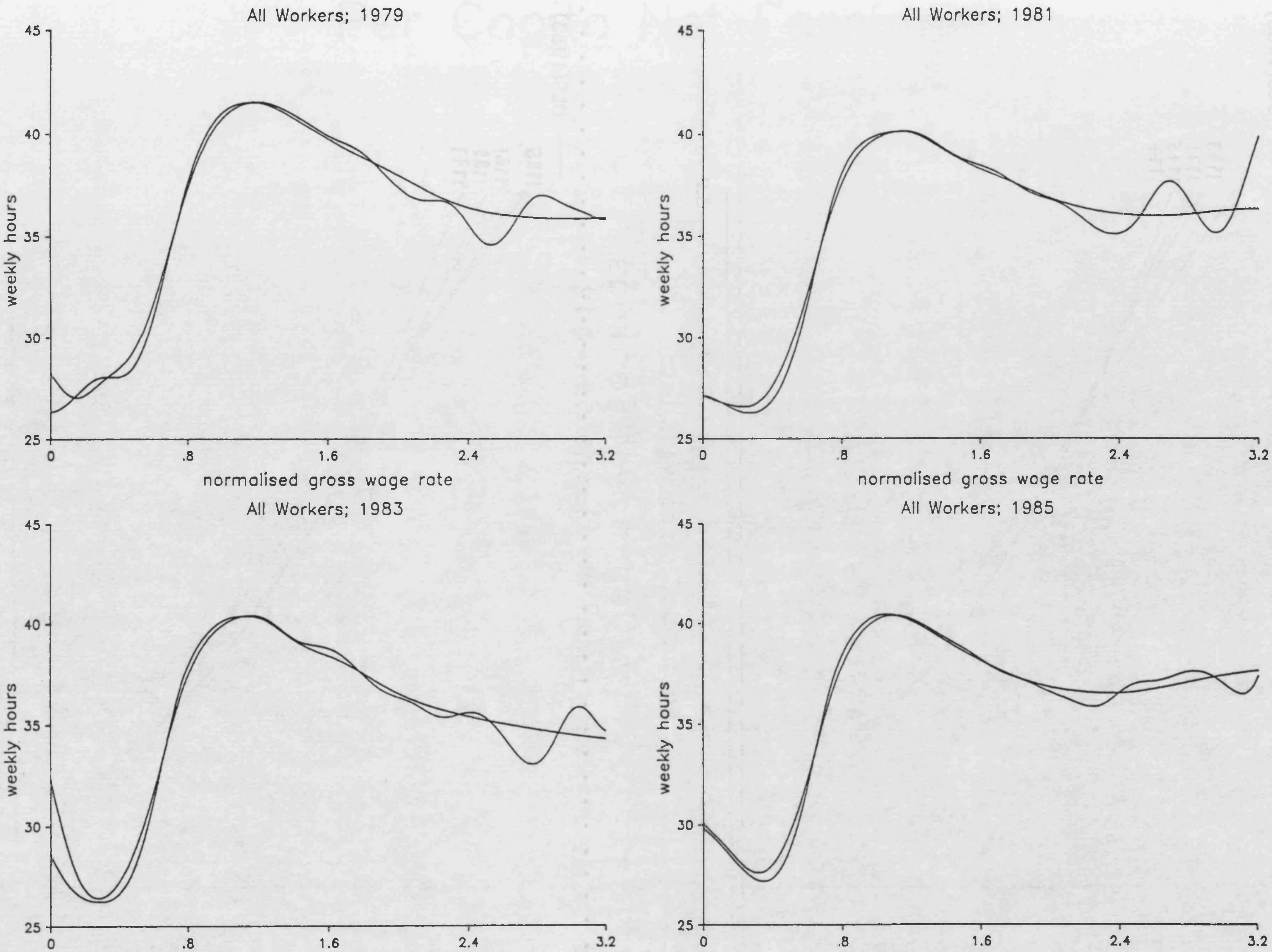


Figure 9c (continued)

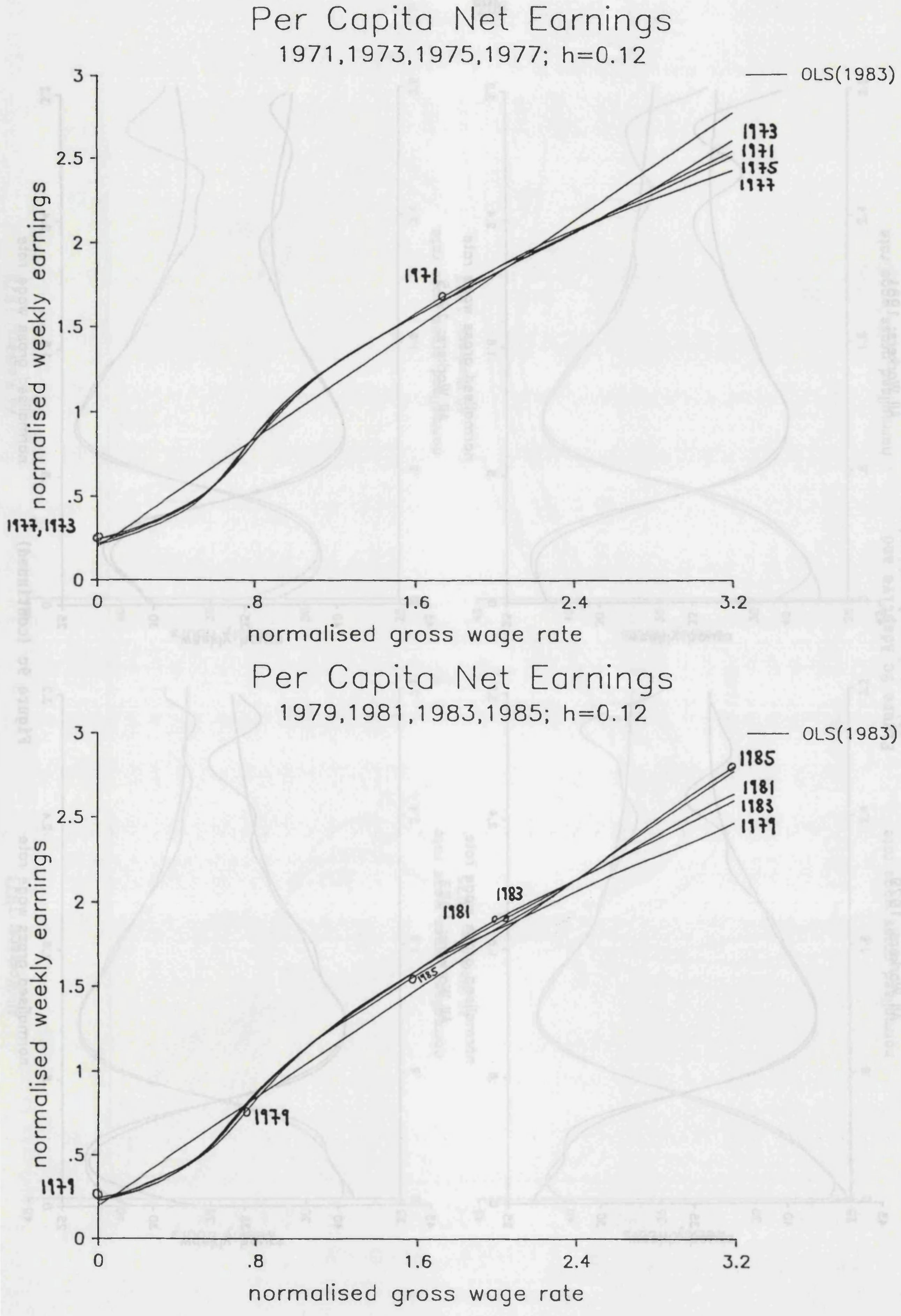


Figure 10a Adaptive kernel estimates

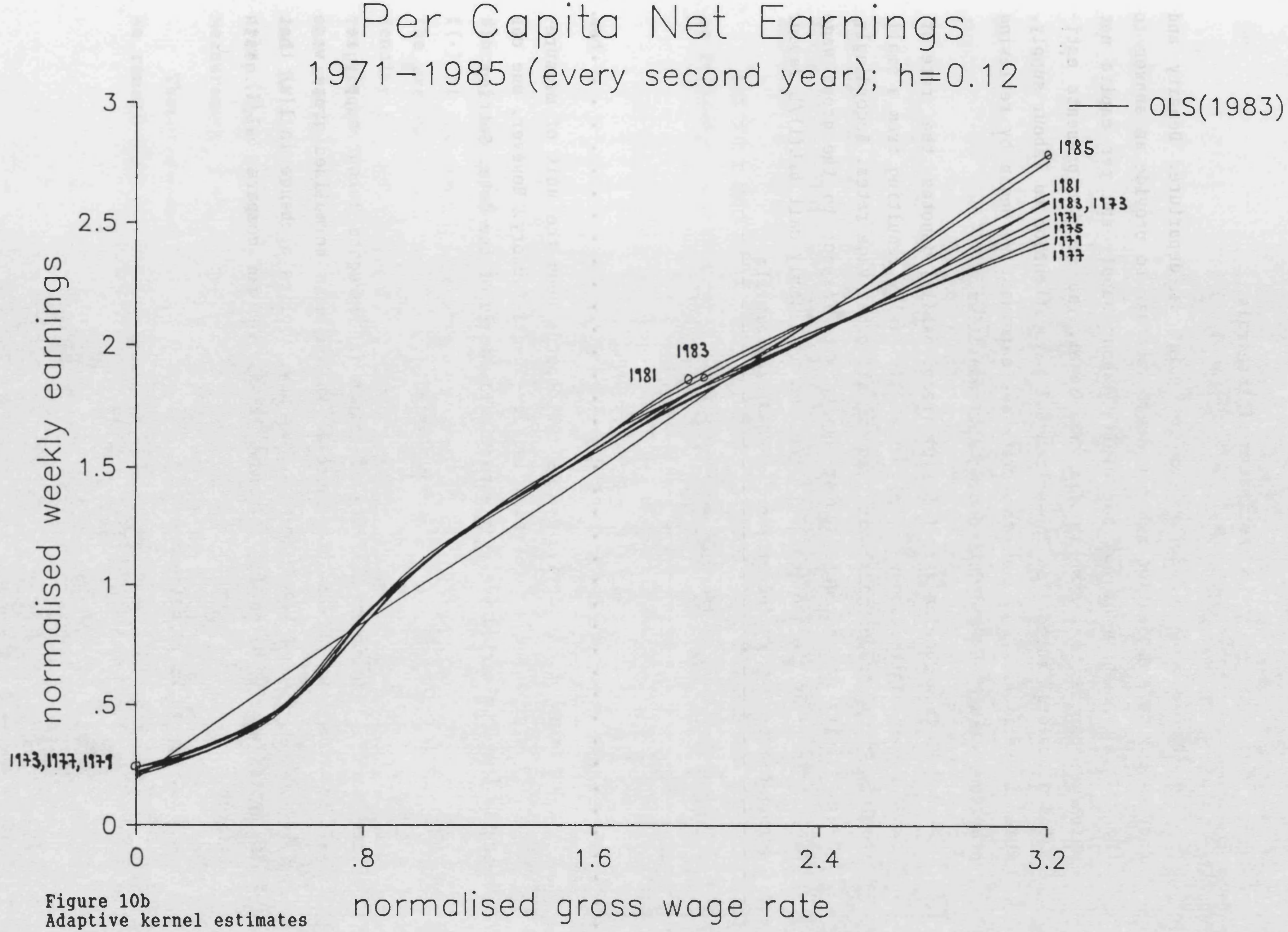


Figure 10b
Adaptive kernel estimates

6. Estimates of Integrals

In this section we return to our point of departure. Density and regression curve estimation are now drawn together to provide an answer to the initial question of how per capita labour supply and per capita net earnings respond to a rise in the wage level. The section presents estimates of several integrals. Subsection 6.1 pays attention to labour supply, while Subsection 6.2 is concerned with net earnings. We begin by reviewing previous definitions and introducing new quantities.

Recall from Section 2 that $\delta_a L(0)$ [resp. $\delta_\alpha L(1)$] denotes the rate of change in per capita labour supply $L = \int l(w)\rho(w)dw$ resulting from a small absolute (resp. proportional) increase in all gross wage rates. Accordingly the elasticity of aggregate labour supply with respect to the gross wage rate is given by the quotient $\delta_\alpha L(1)/L$. We will simply call $\delta_\alpha L(1)/L$ *labour supply elasticity*. By definition of $\delta_a L(0)$ and $\delta_\alpha L(1)$:

$$\delta_a L(0) = \int l'(w)\rho(w)dw$$

and

$$\delta_\alpha L(1) = \int l'(w)w\rho(w)dw.$$

The value of the derivative $\delta_a L(0)$ depends upon the unit of measurement of the gross wage rate which is very unsatisfactory. However, one can easily free the derivative from a particular scale of the data. Setting

$$\mu = \int w\rho(w)dw,$$

then $\delta_a L(0)\mu$ represents the rate of change in aggregate labour supply resulting from a small absolute increase in all mean normalised gross wage rates. Adding to w a fraction of μ, say μdw, implies a change in $l(w)$ that is approximately given by $l'(w)\mu dw$. Thus, we can compare $\delta_\alpha L(1)$ with $\delta_a L(0)\mu$.

Let D_1 and D_2 be defined by

$$D_1 = \int l(w)\rho'(w)dw$$

and

$$D_2 = \int l(w)\cdot(w\rho(w))'dw.$$

Suppose the unknown probability density ρ is concentrated on the interval [a,b] with $\rho(a)=\rho(b)=0$. Then, by partial integration,

$$\delta_a L(0) = -D_1 \text{ and } \delta_\alpha L(1) = -D_2.$$

It is straightforward to obtain inequalities in the case of $\rho(a)>0$ and/or $\rho(b)>0$.

Aggregate net labour income is given by $B = \int b(w)\rho(w)dw$. Substituting $b'(\cdot)$ for $l'(\cdot)$ in the expression for $\delta_a L(0)$ [resp. $\delta_\alpha L(1)$], one obtains the derivative $\delta_a B(0)$ [resp. $\delta_\alpha B(1)$]. The elasticity of net earnings is given by $\delta_\alpha B(1)/B$. The relative rate of change of B resulting from a small absolute increase in w/μ is given by $\delta_a B(0)\mu/B$.

Let $f(\cdot)$ be a differentiable function. We define two further integrals as follows:

$$I_1 = \int \frac{f'(w)w}{f(w)}\rho(w)dw$$

and

$$I_2 = \int \frac{f'(w)}{f(w)}\rho(w)dw.$$

Since $f'(w)$ [resp. $f'(w)w/f(w)$] is the rate of change (resp. elasticity) of $f(\cdot)$ at the wage rate w, I_1 and I_2 are to be interpreted as follows: I_1 is the average elasticity of the function $f(\cdot)$ with respect to the probability density ρ, and I_2 is the average relative rate of change of $f(\cdot)$ with respect to ρ. Hence, setting $f(\cdot)=l(\cdot)$ [resp. $f(\cdot)=b(\cdot)$] then I_1 denotes the average elasticity of the labour supply (resp. net earnings) function, and I_2 denotes the average relative rate of change of the labour supply (resp. net earnings) function. Notice, since $f'(w)$ depends on the unit of measurement of w, I_2 also depends on it.

There are special cases where the above integrals do not depend on ρ. We remark that I_1 (resp. I_2) is independent of ρ if and only if $f(w)=\alpha\cdot w^{\beta}$

[resp. $f(w)=\alpha\cdot\exp\{\beta\cdot w\}$]; $\delta_\alpha L(1)$ does not depend on ρ if and only if $l(w)=\alpha+\beta\cdot\log w$.

In the estimations we used the same scale of the data as the FES. To estimate D_1 and D_2, we need estimates of $l(\cdot)$, $\rho(\cdot)$ and $\rho'(\cdot)$. For the other integrals, estimates of $l(\cdot)$, $l'(\cdot)$ [resp. $b(\cdot)$, $b'(\cdot)$] and $\rho(\cdot)$ are required. The integrals were estimated by substituting kernel smoothers for the unknown functions. We then employed numerical quadrature; as always, the kernel K was the density of the standard normal distribution. The NAG (Mark 12) library routine D01GAF was used to evaluate the integrals. The integrand was specified at the points $x_k=(1/200)\cdot w_{0.99}\cdot k$, $k=0,1,\ldots,200$, where $w_{0.99}$ denotes the ninety-ninth percentile of the empirical distribution of the gross wage rates. Strictly speaking, we did not estimate ρ but the on $[0,w_{0.99}]$ truncated density

$$\tilde{\rho}(w) = \begin{cases} \dfrac{\rho(w)}{\int_0^{w_{0.99}} \rho(w)dw}, & \text{if } w \in [0,w_{0.99}] \\ 0, & \text{otherwise.} \end{cases}$$

In the case of net earnings we compared the results of the kernel estimations with estimates obtained from the least-squares line $b(w)=\alpha+\beta\cdot w$, $w\in R_+$. Of course, linearity of $b(\cdot)$ implies that the probability density ρ can be disregarded; we have $\delta_\alpha B(0)=\beta$ and $\delta_\alpha B(1)=\beta\mu$.

When estimating integrals by the kernel method, the results will depend upon our choice of the smoothing parameter. The estimates which we discuss in the next two subsections are fairly robust. That is, reasonable variations of the smoothing parameter lead only to small changes in the estimates. However, in this study we selected the window width essentially by trial and error. Using screen plots of the kernel smoothers to find parameter values that give enough smoothness without obscuring detail is, of course, an unsatisfactory procedure. Further work is required to obtain more reliable estimates. In particular, confidence intervals for the estimates are needed.

We first draw attention to labour supply.

6.1. Labour Supply

Tables 4-10 on pages 223-226 present the results. The outcomes of the ordinary kernel estimations are given in brackets. The integrals were estimated on each sample separately. The regression curves corresponding to Tables 4-7 are given in Figures 2b, 3b, 4, 6b and 7. The density function was estimated using the same value for the window width as for the corresponding regression curve (in the diagrams of Chapter 2 we used for a density in general a somewhat smaller window width).

We begin by exploring the data of the 1983 FES. Table 4 presents adaptive and ordinary kernel estimates of L, $\delta_\alpha L(1)$, $\delta_a L(0)$, $\delta_\alpha L(1)/L$, I_1 and I_2 for the total population of "all workers" and the eight subpopulations considered throughout this study. Depending on the specific value of the smoothing parameter, we obtain an elasticity of total labour supply of around 0.2. The labour supply elasticity is considerably larger for females than for males. The lowest elasticity occurs in the population of non-manual males, where it is approximately -0.08. The elasticity is largest for manual females: a one per cent increase in the gross wage rate leads to an increase in manual female labour supply of around half a per cent.

When looking at the average elasticities of the labour supply functions (i.e., the I_1 figures), the same picture emerges. The estimates of I_1 and $\delta_\alpha L(1)/L$ differ only very slightly. For example, in the entire population of workers I_1 has the value 0.22. We will see below that excluding part-time workers from the samples drastically reduces the range of labour supply elasticities.

Comparing the ordinary kernel estimates with the adaptive kernel estimates, we see that there is almost no difference in the case of the estimates for per capita labour supply L.[9)] The ordinary kernel method produces, of course, in the upper range of the gross wage rate distribution extremely poor estimates of l'(w) since the regression estimator becomes unstable in this range. Table 4 shows that this has an effect on the estimates of $\delta_\alpha L(1)$, $\delta_a L(0)$, I_1 and I_2. Notice that adaptive kernel estimation leads to a somewhat higher elasticity. But the differences are small, and the estimations confirm what has already been said at the end of

Section 3: The labour supply elasticity is essentially determined by the behaviour of the regression curve in the main body of the gross wage rate distribution.

Since the scale of the gross wage rate data is "tenths of pence per hour", the values of $\delta_a L(0)$ and I_1 are very small. The sample means of the gross wage rates are given in Tables 11-16 on pages 234-237. For instance, $\delta_a L(0)\mu$ assumes on the sample of "all workers" a value of around 12. Notice that for the same group of workers $\delta_\alpha L(1)=7$, approximately. Thus, a small absolute increase in all gross wage rates leads to a larger rise in per capita labour supply than a small proportional increase. This is not surprising and follows immediately from the shape of the aggregate labour supply function (Figure 3). We remark that in all tables the relation between $\delta_a L(0)/L$ and I_2 is the same as that between $\delta_\alpha L(1)/L$ and I_1. We will say more about $\delta_a L(0)\mu$ below.

In Table 5 we see estimates for the population of "all workers 1983" obtained by alternative nonparametric methods. The table is set up as follows. In rows 1, 2, 3 and 7 the regression function was constructed by a spline smoothing of the naive regression estimate (see Subsection 3.1 and Figure 2b); in rows 4, 5 and 6 the regression function was estimated by the adaptive kernel method. Let F be the cumulative distribution function of the unknown gross wage rate distribution, let $f(w) = l(w)$, $l'(w)$, $l'(w)w$, $l'(w)/l(w)$, $l'(w)w/l(w)$, and let $\hat{f}(w)$ denote an estimate of $f(w)$.
Rows 1 and 4 present estimates of

$$\int f(w)\,dF(w),$$

where F was estimated by the empirical cumulative distribution function F_n (see Chapter 2, Section 3). Clearly, by definition of the Stieltjes-integral (see, e.g., Apostol, 1974, Chapter 7, pp. 140-142),

$$\int \hat{f}(w)\,dF_n(w) = \frac{1}{n}\cdot\sum_{i=1}^{n} \hat{f}(w_i),$$

where w_i denotes the i-th observation in the sample of gross wage rates. We averaged over all gross wage rates w_i such that $w_i \leq w_{0.99}$ ($w_{0.99}$ = £10) using a regularly spaced mesh of points $0=x_0<x_1<\ldots<x_{200}=w_{0.99}$; i.e., we set $\hat{f}(w_i)=\hat{f}(t_j)$ for all $w_i\varepsilon]x_{j-1},x_j]$, where $t_j=(x_{j-1}+x_j)/2$ $(j=1,\ldots,200)$. In rows 2, 3, 5 and 6 the unknown probability density was estimated by a spline smoothing of the empirical distribution function; in row 7 a DMPL-estimate of the unknown density was used (see Chapter 2, Subsection 3.2). Rows 3, and 6 provide estimates of $\int l(w)\rho'(w)dw$ and $\int l(w)(\rho(w)w)'dw$.

Table 5 shows that the results are insensitive to the technical detail of the nonparametric estimation procedure. Using the above methods to estimate the elasticity of aggregate labour supply, one obtains values ranging from 0.212 to 0.246, while the adaptive kernel method yields $\delta_\alpha L(1)/L=$ 0.208. Notice that $\delta_\alpha L(1)>-D_2$, and $\delta_a L(0)<-D_1$ when D_1 was estimated by the spline smoothing method. We will return to these inequalities below. From now on we will restrict attention to the kernel method.

The elasticity of total labour supply did not change very much during the years from 1970 to 1985. The results are given in Table 6. The arithmetic mean and the standard deviation of the 16 adaptive kernel estimates of $\delta_\alpha L(1)/L$ are 0.199 and 0.021, respectively. The ordinary kernel method produces somewhat smaller elasticities; their mean value is 0.178 and their standard deviation is 0.017. The average elasticity of the labour supply function is in all years slightly larger than $\delta_\alpha L(1)/L$. We have:

	arithmetic mean	standard deviation
ordinary kernel estimates of I_1:	0.189	0.018
adaptive kernel estimates of I_1:	0.213	0.022

Notice that $\delta_\alpha L(1)/L$ increased in the years 1970-77 from 0.165 to 0.224; I_1 increased from 0.176 to 0.238. During the years 1980-85 $\delta_\alpha L(1)/L$ and I_1 were approximately constant. Clearly, this was to be expected from the shifts in the labour supply curves discussed in Section 5. Looking down the third column of Table 6, we see that $\delta_a L(0)$ is strictly decreasing. The adaptive (resp. ordinary) kernel estimates of $\delta_a L(0)$ have the arithmetic mean 0.0092 (resp. 0.0084) and the standard deviation 0.0053 (resp. 0.0050). Of course, $\delta_a L(0)$ is decreasing since the real value of one unit of w declines throughout the years.

The picture changes when we multiply each estimate of $\delta_a L(0)$ by the mean of the underlying gross wage rate data (the sample means are given in the last column of Table 12 on page 235). The adaptive kernel estimates of $\delta_a L(0)\mu$ have the mean 11.89 and the standard deviation 0.76; the ordinary kernel estimates have the mean 10.75 and the standard deviation 0.61. Looking at every second year between 1970 and 1985, the following picture emerges (the ordinary kernel estimates of $\delta_a L(0)\mu$ are given in brackets):

1971	1973	1975	1977	1979	1981	1983	1985
11.08	10.70	11.83	12.51	12.56	12.33	12.39	11.99
(10.24)	(9.84)	(11.05)	(11.23)	(11.39)	(10.97)	(10.86)	(10.45)

As expected, the corresponding values of $\delta_\alpha L(1)$ are smaller. Using the adaptive kernel method, we obtain for $\delta_\alpha L(1)$ the mean 7.30 and the standard deviation 0.68; in the case of ordinary kernel estimation we have: mean = 6.51 and standard deviation = 0.57. The arithmetic mean and the standard deviation of the adaptive kernel estimates of $\delta_a L(0)\mu/L$ are 0.325 and 0.026, respectively; the ordinary kernel estimates of $\delta_a L(0)\mu/L$ have the mean 0.294 and the standard deviation 0.020. Loosely speaking, this means that observed labour supply rises by approximately 0.3 per cent if we add to all gross wage rates the amount 0.01μ, i.e., one per cent of the mean gross wage rate. This is substantially higher than $\delta_\alpha L(1)/L$.

Turning to the estimates of D_1 and D_2, we see that $-D_2$ is somewhat smaller than $\delta_\alpha L(1)$ in all years. Taking $-D_2/L$ as an estimator for the elasticity of total labour supply, we obtain an average elasticity of 0.158; the standard deviation of the estimates is 0.019. Notice that $\delta_a L(0)$ and $-D_1$ are approximately equal; however $-D_1$ is in 11 of the 16 years very slightly larger than $\delta_a L(0)$.

These discrepancies can be easily explained. Recall that the integrals were estimated on intervals $[0,w_{0.99}]$, where $w_{0.99}$ denotes the ninety-ninth percentile of the underlying empirical gross wage rate distribution. All estimated densities and regression curves assume positive values at $w=0$ and $w=w_{0.99}$. By partial integration, $\delta_\alpha L(1) = l(w_{0.99}) \cdot w_{0.99} \cdot \rho(w_{0.99}) - D_2 > -D_2$. Likewise, we obtain $\delta_a L(0)+D_1 = l(w_{0.99})\rho(w_{0.99})-l(0)\rho(0)$, and the difference on the right-hand side of this equation has in general a negative sign.

Table 7 provides the estimates for full-time workers (Figures 4 and 7). Only in the case of manual workers are the signs of $\delta_a L(0)$ and $\delta_\alpha L(1)$ positive. However, the elasticity of full-time manual labour supply has the extremely small value of 0.01; and $\delta_a L(0)\mu/L=0.015$, approximately. In the remaining groups of full-time workers the values of $\delta_\alpha L(1)/L$ are ranging from -0.098 (for non-manual males) to -0.014 (for manual females); in the total population of "all full-time workers 1983" the elasticity is -0.02.

Thus, the exclusion of part-time workers from the samples has impressively narrowed down the range of labour supply elasticities. From the point of view of $\delta_\alpha L(1)/L$, there is no difference between female and male labour supply any more. The picture does not change when looking at $\delta_a L(0)\mu/L$. For instance, in the total population of full-time workers we have: $\delta_a L(0)\mu/L|_{adapt.\ kernel} = -0.017$ and $\delta_a L(0)\mu/L|_{ord.\ kernel} = -0.026$.

We give two examples showing the effect of varying the window width. In rows 6 (manual males) and 8 (manuals) of Table 7 the density and the regression curve were estimated using h=390. Decreasing h by 30 per cent produces regression curves which exhibit in the entire range of ρ more local variability. However, the effect on $\delta_\alpha L(1)/L$ is approximately zero. We obtain the following elasticties: "All Manuals" = 0.013 (0.007) and "Manual Males" = -0.031 (-0.032), where the figures in brackets are the ordinary kernel estimates of $\delta_\alpha L(1)/L$.

In the remainder of this subsection we pay attention to male labour supply. In the case of male manual workers the exclusion of part-time workers implied a change in the sign of the labour supply elasticity. However, since around 94 per cent of the males work full-time, the differences between the figures in Tables 4 and 7 are very small. Observed male labour supply is extremely insensitive to small variations in the gross wage rate. Depending on the particular sample considered, the rate of change in male labour supply resulting from a 1 per cent increase in the gross wage rate ranges from -0.1% (for "full-time non-manual males") to +0.05% (for "all manual males"). We estimated the response of male labour supply to a rise in the gross wage rate for each odd numbered year from 1971 to 1985 on the sample of "all male workers" and on the subsamples of "manual males" and "non-manual males" (the samples contain part-time workers). As we will see, the estimates are remarkably stable.

Table 8 presents the estimates for the subgroup "male workers". Reading down the fourth column of the table, we see that the elasticity of male labour supply is sligtly negative in all years. Notice that $|\delta_\alpha L(1)/L|$ is increasing over the years 1971-77 and decreasing in the remaining years. The arithmetic mean and the standard deviation of the eight adaptive kernel estimates of the labour supply elasticity are -0.042 and 0.011, respectively. The ordinary kernel estimates have the mean -0.040 and the standard deviation 0.010. Observe that in all years $\delta_\alpha L(1)=I_1$, approximately. As always, $\delta_a L(0)$ and I_2 are very small. Multiplying $\delta_a L(0)$ by the mean gross wage rate yields (the ordinary kernel estimates are given in brackets):

1971	1973	1975	1977	1979	1981	1983	1985
+1.47	+0.60	+0.43	-0.79	-0.06	-0.03	+1.00	+1.15
(+2.01)	(+0.81)	(+1.13)	(-0.14)	(+0.39)	(+0.66)	(+1.35)	(+1.46)

Observe that only in 1977 both the ordinary and the adaptive kernel estimate of $\delta_a L(0)\mu$ have a negative sign; $|\delta_a L(0)\mu|$ is somewhat smaller than $|\delta_\alpha L(1)|$. We have:

	arithmetic mean	standard deviation
ordinary kernel estimates of $\delta_a L(0)\mu$:	+0.96	0.63
adaptive kernel estimates of $\delta_a L(0)\mu$:	+0.47	0.70
ordinary kernel estimates of $\delta_\alpha L(1)$:	-1.70	0.45
adaptive kernel estimates of $\delta_\alpha L(1)$:	-1.79	0.47
ordinary kernel estimates of $\delta_a L(0)\mu/L$:	+0.02	0.015
adaptive kernel estimates of $\delta_a L(0)\mu/L$:	-0.01	0.016

Turning to manual male labour supply (Table 9), we see that apart from 1977 all estimates of $\delta_\alpha L(1)/L$ and I_1 have a positive sign. The labour supply elasticity is larger during the years from 1979 to 1985 than in the first part of the period; $\delta_\alpha L(1)/L$ is decreasing in the years 1971-77 and increasing in the years 1977-85. The average elasticity of the labour supply function, which is also first decreasing and then increasing, is in all years slightly larger than $\delta_\alpha L(1)/L$. The arithmetic means and the standard deviations of the estimates of $\delta_\alpha L(1)/L$ and I_1 are as follows:

	arithmetic mean	standard deviation
ordinary kernel estimates of $\delta_\alpha L(1)/L$:	0.017	0.017
adaptive kernel estimates of $\delta_\alpha L(1)/L$:	0.020	0.018
ordinary kernel estimates of I_1:	0.021	0.018
adaptive kernel estimates of I_1:	0.024	0.019

$\delta_a L(0)$ and I_2 are positive in all years. Taking the sample means from Table 15 (page 237), we obtain the following time-series for $\delta_a L(0)\mu$ (the figures in brackets are the ordinary kernel estimates):

1971	1973	1975	1977	1979	1981	1983	1985
4.14	2.63	2.78	2.20	3.35	3.07	4.48	5.00
(4.34)	(2.57)	(3.17)	(2.45)	(3.52)	(3.39)	(4.45)	(5.07)

We have:

	arithmetic mean	standard deviation
ordinary kernel estimates of $\delta_a L(0)\mu$:	3.62	0.87
adaptive kernel estimates of $\delta_a L(0)\mu$:	3.46	0.92
ordinary kernel estimates of $\delta_\alpha L(1)$:	0.74	0.74
adaptive kernel estimates of $\delta_\alpha L(1)$:	0.88	0.78
ordinary kernel estimates of $\delta_a L(0)\mu/L$:	0.084	0.021
adaptive kernel estimates of $\delta_a L(0)\mu/L$:	0.080	0.022

Since the labour supply elasticity for the group "male workers" has a negative sign and that for "male manual workers" has a positive sign, the elasticity of non-manual male labour supply must be negative. Table 10 provides the estimates. (Strictly speaking, one does not have to carry out separately the estimations on the annual samples of "non-manual male workers". It is straightforward to obtain estimates for this group of workers from Tables 8 and 9 and the sample sizes given in the Appendix.)
We have:

	arithmetic mean	standard deviation
ordinary kernel estimates of $\delta_\alpha L(1)/L$:	-0.064	0.015
adaptive kernel estimates of $\delta_\alpha L(1)/L$:	-0.071	0.016
ordinary kernel estimates of I_1:	-0.065	0.016
adaptive kernel estimates of I_1:	-0.072	0.017

The values for $\delta_a L(0)\mu$ are as follows (the ordinary kernel estimates are shown in brackets):

1971	1973	1975	1977	1979	1981	1983	1985
-1.77	-1.05	-1.42	-3.35	-2.90	-3.16	-2.92	-2.79
(-1.08)	(-0.72)	(-0.63)	(-2.62)	(-2.63)	(-2.37)	(-2.19)	(-2.27)

Again we have $\delta_\alpha L(1) < \delta_a L(0)\mu$ in all years. From Table 7 and the estimates of $\delta_a L(0)\mu$ we obtain:

	arithmetic mean	standard deviation
ordinary kernel estimates of $\delta_a L(0)\mu$:	-1.81	0.80
adaptive kernel estimates of $\delta_a L(0)\mu$:	-2.42	0.82
ordinary kernel estimates of $\delta_\alpha L(1)$:	-2.60	0.62
adaptive kernel estimates of $\delta_\alpha L(1)$:	-2.89	0.66
ordinary kernel estimates of $\delta_a L(0)\mu/L$:	-0.045	0.020
adaptive kernel estimates of $\delta_a L(0)\mu/L$:	-0.060	0.020

Summing up, it is clear from the data analysis that labour supply is not very responsive to a change in the gross wage rate. In particular, the elasticity of total labour supply of around 0.2 can be explained by the high proportion of part-time workers in the labour force. It is interesting to observe that excluding part-time workers from the 1983 FES data produces an elasticity close to zero. A general feature of the data is that $\delta_a L(0)\mu/L$ is somewhat larger than $\delta_\alpha L(1)/L$. However, regardless of whether we consider $\delta_a L(0)\mu/L$ or $\delta_\alpha L(1)/L$, the qualitative picture is always the same. Observed labour supply is inelastic, and when looking at full-time workers or populations which consist mainly of full-time workers, it is even extremely insensitive to small variations in the gross wage rate.

It would be interesting to decompose the actual observed time-series (l_t, ρ_t) into (l_0, ρ_t) and (l_t, ρ_0), t=0 denoting any base year, and to quantify the effects of changes in l_t (resp. ρ_t) on the elasticity.

We now turn to per capita net earnings. Notice that $I_{1,\text{net earnings}}$ is equal to $1 + I_{1,\text{hours}}$ if the tax function is linear. A brief description of the tax function was given at the end of Subsection 3.1 of Chapter 2.

Table 4 Labour Supply

All workers and subgroups, 1983; adaptive and ordinary kernel estimates

Ordinary kernel estimates in parentheses

Sample	L	$\delta_a L(1)$	$\delta_a L(0)$	$\frac{\delta_a L(1)}{L}$	I_1	I_2
All Workers	35.63 (35.62)	7.40 (6.56)	0.389E-2 (0.341E-2)	0.208 (0.184)	0.224 (0.197)	0.119E-3 (0.103E-3)
Females	28.01 (28.02)	7.75 (6.57)	0.413E-2 (0.330E-2)	0.277 (0.234)	0.285 (0.238)	0.153E-3 (0.121E-3)
Manual Females	24.56 (24.58)	12.75 (11.73)	0.638E-2 (0.548E-2)	0.519 (0.477)	0.507 (0.461)	0.257E-3 (0.219E-3)
Non-Manual Females	29.86 (29.85)	3.85 (2.75)	0.236E-2 (0.181E-2)	0.129 (0.092)	0.131 (0.087)	0.817E-4 (0.616E-4)
Males	41.69 (41.65)	-1.37 (-1.34)	0.270E-3 (0.362E-3)	-0.033 (-0.032)	-0.033 (-0.032)	0.742E-5 (0.992E-5
Manual Males	42.40 (42.38)	2.07 (1.90)	0.150E-2 (0.149E-2)	0.049 (0.045)	0.053 (0.050)	0.385E-4 (0.388E-4)
Non-Manual Males	40.68 (40.66)	-3.51 (-3.21)	-0.629E-3 (-0.471E-3)	-0.086 (-0.079)	-0.087 (-0.080)	-0.154E-4 (-0.115E-4)
Manuals	36.46 (36.45)	13.37 (12.57)	0.645E-2 (0.597E-2)	0.367 (0.345)	0.394 (0.367)	0.193E-3 (0.177E-3)
Non-Manuals	34.83 (34.83)	4.14 (3.62)	0.216E-2 (0.185E-2)	0.119 (0.104)	0.126 (0.109)	0.659E-4 (0.560E-4)

Table 5 Alternative Nonparametric Estimation Methods

All workers 1983: labour supply

Method	L	$\delta_a L(1)$	$\delta_a L(0)$	$\frac{\delta_a L(1)}{L}$	I_1	I_2	$-D_1$	$-D_2$
emp. cdf l: spline	35.45	8.63	0.474E-2	0.243	0.270	0.151E-3		
ρ: spline l: spline	35.51	8.75	0.472E-2	0.246	0.272	0.149E-3		
ρ: spline l: spline							0.492E-2	5.47
emp. cdf l: kernel	35.39	7.52	0.399E-2	0.212	0.229	0.123E-3		
ρ: spline l: kernel	35.62	7.56	0.399E-2	0.212	0.230	0.122E-3		
ρ: spline l: kernel							0.417E-2	4.34
ρ: DMPL l: spline	35.50	8.61	0.471E-2	0.243	0.268	0.149E-3		

Table 6 All Workers: Labour Supply
Adaptive and ordinary kernel estimates[1]; global window width:
h = 0.12·(mean gross wage rate)

Sample	L	$\delta_w L(1)$	$\delta_a L(0)$	$\frac{\delta_w L(1)}{L}$	I_1	I_2	$-D_1$	$-D_2$
1970	38.34 (38.33)	6.32 (5.87)	0.199E-1 (0.186E-1)	0.165 (0.153)	0.176 (0.163)	0.555E-3 (0.518E-3)	0.201E-1	4.73
1971	37.66 (37.64)	6.25 (5.68)	0.185E-1 (0.171E-1)	0.166 (0.151)	0.179 (0.161)	0.531E-3 (0.490E-3)	0.189E-1	4.89
1972	37.62 (37.61)	5.80 (5.22)	0.152E-1 (0.140E-1)	0.154 (0.139)	0.165 (0.148)	0.434E-3 (0.402E-3)	0.156E-1	4.43
1973	37.24 (37.23)	6.53 (6.01)	0.138E-1 (0.127E-1)	0.175 (0.161)	0.187 (0.171)	0.395E-3 (0.359E-3)	0.140E-1	5.08
1974	37.13 (37.09)	7.77 (7.18)	0.136E-1 (0.129E-1)	0.209 (0.194)	0.224 (0.206)	0.399E-3 (0.377E-3)	0.136E-1	6.26
1975	37.04 (37.02)	7.52 (6.90)	0.102E-1 (0.953E-2)	0.203 (0.186)	0.216 (0.198)	0.298E-3 (0.279E-3)	0.101E-1	5.82
1976	36.62 (36.59)	7.78 (6.89)	0.874E-2 (0.789E-2)	0.213 (0.188)	0.223 (0.197)	0.253E-3 (0.229E-3)	0.886E-2	6.51
1977	36.72 (36.70)	8.22 (7.23)	0.839E-2 (0.753E-2)	0.224 (0.197)	0.238 (0.208)	0.245E-3 (0.219E-3)	0.833E-2	6.43
1978	36.70 (36.67)	7.59 (6.54)	0.706E-2 (0.625E-2)	0.207 (0.178)	0.220 (0.188)	0.206E-3 (0.182E-3)	0.709E-2	6.02
1979	36.86 (36.83)	8.20 (7.35)	0.657E-2 (0.596E-2)	0.222 (0.199)	0.237 (0.212)	0.192E-3 (0.174E-3)	0.653E-2	6.45
1980	36.29 (36.27)	7.54 (6.67)	0.519E-2 (0.466E-2)	0.208 (0.184)	0.222 (0.196)	0.154E-3 (0.138E-3)	0.520E-2	5.96
1981	35.87 (35.85)	7.41 (6.46)	0.454E-2 (0.404E-2)	0.206 (0.180)	0.223 (0.193)	0.138E-3 (0.122E-3)	0.455E-2	5.76
1982	35.58 (35.55)	7.37 (6.48)	0.442E-2 (0.392E-2)	0.207 (0.182)	0.227 (0.198)	0.137E-3 (0.121E-3)	0.451E-2	6.31
1983	35.63 (35.62)	7.40 (6.56)	0.389E-2 (0.341E-2)	0.208 (0.184)	0.224 (0.197)	0.119E-3 (0.103E-3)	0.398E-2	6.10
1984	35.50 (35.48)	7.80 (6.80)	0.380E-2 (0.335E-2)	0.220 (0.192)	0.237 (0.205)	0.116E-3 (0.101E-3)	0.379E-2	5.82
1985	35.82 (35.80)	7.25 (6.37)	0.327E-2 (0.285E-2)	0.202 (0.178)	0.217 (0.189)	0.985E-4 (0.853E-4)	0.336E-2	5.85

1) Ordinary kernel estimates in parentheses.

Table 7 Full-Time Workers, 1983: Labour Supply
Adaptive and ordinary kernel estimates[1)]

Sample	L	$\delta_a L(1)$	$\delta_a L(0)$	$\frac{\delta_a L(1)}{L}$	I_1	I_2
All Workers (Full-Time)	41.63 (41.64)	-0.91 (-1.06)	-0.208E-3 (-0.313E-3)	-0.022 (-0.025)	-0.022 (-0.026)	-0.502E-5 (-0.744E-5)
Females	38.38 (38.40)	-1.74 (-1.64)	-0.969E-3 (-0.107E-2)	-0.045 (-0.043)	-0.045 (-0.042)	-0.247E-4 (-0.266E-4)
Manual Females	39.52 (39.55)	-0.57 (-0.58)	-0.676E-3 (-0.766E-3)	-0.014 (-0.015)	-0.014 (-0.014)	-0.166E-4 (-0.186E-4)
Non-Manual Females	37.97 (37.99)	-1.14 (-1.37)	-0.780E-3 (-0.904E-3)	-0.030 (-0.036)	-0.038 (-0.036)	-0.202E-4 (-0.224E-4)
Males	43.02 (43.02)	-3.33 (-3.19)	-0.942E-3 (-0.919E-3)	-0.077 (-0.074)	-0.078 (-0.075)	-0.219E-4 (-0.213E-4)
Manual Males	43.82 (43.83)	-1.26 (-1.31)	-0.409E-3 (-0.442E-3)	-0.029 (-0.030)	-0.029 (-0.030)	-0.934E-5 (-0.101E-4)
Non-Manual Males	41.94 (41.95)	-4.12 (-3.86)	-0.114E-2 (-0.109E-2)	-0.098 (-0.092)	-0.098 (-0.092)	-0.268E-4 (-0.256E-4)
Manuals	43.11 (43.12)	0.52 (0.26)	0.268E-3 (0.135E-3)	0.012 (0.006)	0.012 (0.006)	0.630E-5 (0.319E-5)
Non-Manuals	40.28 (40.30)	-0.77 (-0.85)	-0.284E-3 (-0.373E-3)	-0.019 (-0.021)	-0.019 (-0.021)	-0.694E-5 (-0.895E-5)

Table 8 Male Workers: Labour Supply
Adaptive and ordinary kernel estimates[1)]; global window width:
h = 0.12·(mean gross wage rate)

YEAR	L	$\delta_a L(1)$	$\delta_a L(0)$	$\frac{\delta_a L(1)}{L}$	I_1	I_2
1971	42.62 (42.57)	-1.47 (-1.40)	0.208E-2 (0.284E-2)	-0.034 (-0.033)	-0.033 (-0.031)	0.576E-4 (0.804E-4)
1973	42.84 (42.82)	-1.49 (-1.45)	0.650E-3 (0.873E-3)	-0.035 (-0.034)	-0.035 (-0.034)	0.171E-4 (0.227E-4)
1975	42.79 (42.76)	-1.83 (-1.72)	0.319E-3 (0.838E-3)	-0.043 (-0.040)	-0.043 (-0.039)	0.967E-5 (0.245E-4)
1977	42.66 (42.62)	-2.67 (-2.54)	-0.462E-3 (-0.794E-4)	-0.063 (-0.060)	-0.064 (-0.060)	-0.105E-4 (-0.660E-7)
1979	42.95 (42.91)	-2.19 (-2.16)	-0.255E-4 (0.174E-3)	-0.051 (-0.050)	-0.052 (-0.051)	-0.901E-7 (0.506E-5)
1981	41.76 (41.72)	-2.11 (-1.83)	-0.931E-5 (0.208E-3)	-0.051 (-0.044)	-0.051 (-0.043)	0.628E-6 (0.658E-5)
1983	41.69 (41.65)	-1.37 (-1.34)	0.270E-3 (0.362E-3)	-0.033 (-0.032)	-0.033 (-0.032)	0.742E-5 (0.992E-5)
1985	41.99 (41.97)	-1.18 (-1.15)	0.267E-3 (0.337E-3)	-0.028 (-0.027)	-0.028 (-0.027)	0.698E-5 (0.890E-5)

Table 9 Male Manual Workers: Labour Supply

Adaptive and ordinary kernel estimates[1]; global window width: h = 0.09·(mean gross wage rate)

YEAR	L	$\delta_w L(1)$	$\delta_a L(0)$	$\frac{\delta_w L(1)}{L}$	I_1	I_2
1971	43.60 (43.56)	0.87 (0.72)	0.693E-2 (0.726E-2)	0.020 (0.017)	0.026 (0.023)	0.182E-3 (0.196E-3)
1973	43.91 (43.90)	0.33 (0.16)	0.339E-2 (0.331E-2)	0.008 (0.004)	0.010 (0.006)	0.831E-4 (0.817E-4)
1975	43.98 (43.94)	0.26 (0.19)	0.235E-2 (0.268E-2)	0.006 (0.004)	0.009 (0.008)	0.600E-4 (0.712E-4)
1977	43.74 (43.71)	-0.13 (-0.25)	0.148E-2 (0.165E-2)	-0.003 (-0.006)	-0.001 (-0.003)	0.369E-4 (0.438E-4)
1979	44.12 (44.09)	0.86 (0.77)	0.175E-2 (0.184E-2)	0.019 (0.017)	0.023 (0.021)	0.436E-4 (0.467E-4)
1981	42.69 (42.65)	0.57 (0.50)	0.117E-2 (0.129E-2)	0.013 (0.012)	0.017 (0.016)	0.306E-4 (0.348E-4)
1983	42.40 (42.38)	2.07 (1.90)	0.150E-2 (0.149E-2)	0.049 (0.045)	0.053 (0.050)	0.385E-4 (0.388E-4)
1985	42.90 (42.86)	2.17 (1.90)	0.144E-2 (0.146E-2)	0.051 (0.044)	0.055 (0.049)	0.364E-4 (0.376E-4)

Table 10 Male Non-Manual Workers: Labour Supply

Adaptive and ordinary kernel estimates[1]; global window width: h = 0.14·(mean gross wage rate)

YEAR	L	$\delta_w L(1)$	$\delta_a L(0)$	$\frac{\delta_w L(1)}{L}$	I_1	I_2
1971	40.53 (40.49)	-2.42 (-2.06)	-0.189E-2 (-0.115E-2)	-0.060 (-0.051)	-0.060 (-0.051)	-0.462E-4 (-0.273E-4)
1973	40.79 (40.77)	-1.68 (-1.61)	-0.871E-3 (-0.602E-3)	-0.041 (-0.039)	-0.041 (-0.040)	-0.212E-4 (-0.146E-4)
1975	40.58 (40.54)	-2.22 (-2.04)	-0.855E-3 (-0.379E-3)	-0.055 (-0.050)	-0.055 (-0.050)	-0.208E-4 (-0.786E-5)
1977	40.57 (40.53)	-3.53 (-3.33)	-0.157E-2 (-0.123E-2)	-0.087 (-0.082)	-0.088 (-0.083)	-0.388E-4 (-0.301E-4)
1979	40.84 (40.84)	-3.46 (-3.22)	-0.106E-2 (-0.964E-3)	-0.085 (-0.079)	-0.086 (-0.080)	-0.263E-4 (-0.240E-4)
1981	40.30 (40.29)	-3.40 (-2.97)	-0.793E-3 (-0.594E-3)	-0.084 (-0.074)	-0.084 (-0.074)	-0.195E-4 (-0.146E-4)
1983	40.68 (40.66)	-3.51 (-3.21)	-0.629E-3 (-0.471E-3)	-0.086 (-0.079)	-0.087 (-0.080)	-0.154E-4 (-0.115E-4)
1985	40.69 (40.70)	-2.89 (-2.39)	-0.509E-3 (-0.413E-3)	-0.071 (-0.059)	-0.071 (-0.059)	-0.125E-4 (-0.101E-4)

1) Ordinary kernel estimates in parentheses.

6.2. Net Earnings

The results are summerised in Tables 11-16 on pages 234-237. The tables present adaptive kernel estimates of B, $\delta_a B(0)$, $\delta_\alpha B(1)$, $\delta_\alpha B(1)/B$, I_1 and I_2 as well as estimates obtained from the least-squares line b(w)=a+bw; the standard errors of a and b are shown in brackets. Denoting the arithmetic mean of the gross wage rate data by $\bar{W}$, then the OLS estimates of $\delta_a B(0)$, $\delta_\alpha B(1)$ and $\delta_\alpha B(1)/B$ are given by b, $b\bar{W}$ and $b\bar{W}/a+b\bar{W}$, respectively. Clearly, by definition of a and b, $a+b\bar{W}$ is just the arithmetic mean of the net earnings data.

The tables are set up as follows: columns 1-6 contain the adaptive kernel estimates; columns 7-8 contain the OLS estimates, and the sample means $\bar{W}$ are given in the last column. In the estimations the same values for the window width were used as in the corresponding tables of the previous subsection. The regression curves corresponding to Tables 11-13 are given in Figures 5, 8 and 10. We begin again with the data of the 1983 FES.

Looking down the fourth column of Table 11, we see the adaptive kernel estimates of the net earnings elasticity for the population of "all workers 1983" and the eight subpopulations. The values of $\delta_\alpha B(1)/B$ are ranging from 0.764 to 1.181; in the total population of workers the elasticity is 0.939. The qualitative picture is the same as in the case of labour supply:

(1) the elasticity is substantially larger for females than for males;

(2) in both populations $\delta_\alpha B(1)/B$ is larger in the subgroup of manuals than in that of non-manuals;

(3) the net earnings elasticity of "all manual workers" is larger than that of "all non-manual workers" [note that (2) does not necessarily imply (3)].

Clearly, this was to be expected. However, when comparing the tenth column of the table with the fourth, we observe considerable differences between the OLS and the adaptive kernel estimates of $\delta_\alpha B(1)/B$.

With two exceptions (namely "manual males" and "all manuals") the least-squares approach leads to a smaller elasticity. For example, with this method we obtain an overall net earnings elasticity of 0.806. In the case of "female workers" the reduction in $\delta_\alpha B(1)/B$ is so large that the qualitative picture of the data is even changed. Under the assumption that the regression functions can be approximated by straight lines, the net earnings of males are slightly more elastic than those of females. The elasticity is now lowest in the subgroup of "non-manual female workers".

Table 12 displays the time-series of estimation results for the population of "all workers". The adaptive kernel method gives us a net earnings elasticity of around 0.93 in all years. Notice that $\delta_\alpha B(1)/B$ is slightly increasing (resp. decreasing) in the years 1970-80 (resp. 1980-85). As in the case of labour supply, the average elasticity I_1 of the regression function is somewhat larger than $\delta_\alpha B(1)/B$.

There are substantial differences between the OLS and the kernel estimates. Firstly, OLS estimation leads to a lower elasticity in all years. Secondly, the OLS estimates exhibit significant changes over the years: while the kernel estimates deviate on average by only around 1 per cent from its mean value, the OLS estimations produce an average deviation from the mean of some 14 per cent.

As expected from what has been said at the end of Section 3, the slope, b, of the least-squares line is a fairly poor estimator for the average value of b'(w) with respect to the gross wage rate distribution. Notice that the kernel estimates of $\delta_a B(0)$ are (strictly) decreasing from 1970 to 1976, increasing during the years from 1976 to 1979 and again decreasing from then onwards. Comparison of the value for 1970 with that for 1985 shows a decrease of around 16 per cent. The values of b are neither strictly decreasing during the years 1970-76 and 1979-85, nor are they strictly increasing from 1976 to 1979; furthermore b is smaller in 1970 than in 1985.

The means and the standard deviations of the estimates are:

	arithmetic mean	standard deviation
adaptive kernel estimates of $\delta_\alpha B(1)/B$:	0.936	0.012
least-squares estimates of $\delta_\alpha B(1)/B$:	0.789	0.113
adaptive kernel estimates of I_1:	0.979	0.014
adaptive kernel estimates of $\delta_a B(0)$:	27.63	1.49
least-squares estimates of $\delta_a B(0)$:	21.83	3.57

The standard error of b is small and approximately constant throughout the years. The mean and the standard deviation of the 16 standard errors are 0.20 and 0.02, respectively. The standard error of a depends upon the scale of the net earnings data. Since the variance of the data steadily increases from 1970 to 1985, the standard errror of a is also increasing (from 168 in 1970 to 901 in 1985).

If $b(w) = a + b \cdot w$, then the change in B resulting from an increase in w does not depend upon the particular way in which the gross wage rates are increased in the population but only on the average increase per person. The kernel estimations, however, show that a uniform absolute increase in all gross wage rates implies a somewhat larger change in B than a uniform proportional increase. Looking at each second year between 1970 and 1985, we obtain the following values for $\delta_a B(0)\mu/B$:

1971	1973	1975	1977	1979	1981	1983	1985
1.020	1.011	1.012	1.029	1.031	1.037	1.029	1.013

The sixteen kernel estimates of $\delta_a B(0)\mu/B$ for the years 1970-85 have the mean 1.022 and the standard deviation 0.010.

Table 13 presents the results for full-time workers. Let us begin with the adaptive kernel estimates: Excluding part-time workers leads to a lower elasticity in eight of the nine samples. The net earnings elasticities range from 0.704 (for manual females) to 0.840 (for non-manuals); on the entire sample of "all full-time workers 1983" we obtain an elasticity of 0.804. In the case of non-manual males the exclusion of part-time workers leads to a very slight increase in $\delta_\alpha B(1)/B$ of 0.005. The reductions in $\delta_\alpha B(1)/B$ for the other groups are ranging from 0.017 (for males) to 0.477 (for manual females). Clearly, for male workers the differences between the figures in Tables 11 and 13 are small since most of them work full-time.

The OLS estimates give us a different picture of the data. The least-squares estimates of $\delta_\alpha B(1)/B$ are larger than the adaptive kernel estimates (the only exception is the subgroup "full-time non-manual workers"). Excluding part-time workers from the samples implies now an increase in the net earnings elasticity for six of the nine populations; in the case of manual workers the OLS-elasticity is in Table 13 smaller than in Table 11 (for females as well as for males).

As in Table 11, the OLS estimates imply another ranking of the elasticities. When using the kernel method, $\delta_\alpha B(1)/B$ is somewhat larger in the population of "full-time male workers" than in that of "full-time female workers", and in both populations the elasticity is somewhat larger in the subgroup of non-manuals than in that of manuals [this is interesting, since $\delta_\alpha L(1)/L$ is slightly larger for females than for males, and it is also slightly larger for manual males (resp. manual females) than for non-manual males (resp. non-manual females); see Table 7]. According to the OLS estimations, the net earnings of non-manual females are also more elastic than those of manual females. But the OLS estimates reverse the other relations. In particular, the elasticity is now larger for "all full-time manual workers" than for "all full-time non-manual workers". Thus, the ranking of the net earnings elasticities implied by the linear regression model agrees very well with that of the corresponding labour supply elsticities displayed in Table 7.

We now turn to the net earnings elasticity of male workers in the years 1970-85. As in the case of labour supply, we performed the estimations for each odd numbered year on the sample of "all male workers" and on the subsamples of "manual males" and "non-manual males" (the samples contain part-time workers).

The results for the entire population of male workers are reported in Table 14. The kernel estimates of $\delta_\alpha B(1)/B$ and I_1 differ in all years only very slightly. The OLS estimate of the net earnings elasticity is in 1971, 1975, 1977 and 1979 larger than the corresponding kernel estimate. Recall from the preceding subsection that $\delta_\alpha L(1)/L$ is decreasing over the years 1971-77 and increasing between 1977 and 1985 (Table 8). Reading down the fourth column of Table 14, we see that the kernel estimates of the net earnings elasticity - as well as the estimates of I_1 - are also first decreasing and then increasing. The OLS estimates of $\delta_\alpha B(1)/B$, however, are decreasing from 1979 to 1985.

Both the least-squares estimates and the kernel estimates of $\delta_a B(0)$ are decreasing from 1971 to 1977, increasing between 1977 and 1979 and again decreasing from then onwards. Notice that in each year the kernel estimate of $\delta_a B(0)$ is somewhat larger than the corresponding OLS estimate. The kernel estimates of $\delta_a B(0)\mu/B$ are as follows:

1971	1973	1975	1977	1979	1981	1983	1985
0.858	0.837	0.805	0.780	0.809	0.825	0.841	0.826

The means and the standard deviations of the estimates are:

	arithmetic mean	standard deviation
adaptive kernel estimates of $\delta_\alpha B(1)/B$:	0.765	0.017
least-squares estimates of $\delta_\alpha B(1)/B$:	0.747	0.047
adaptive kernel estimates of I_1:	0.764	0.018
adaptive kernel estimates of $\delta_a B(0)$:	24.22	1.46
least-squares estimates of $\delta_a B(0)$:	22.48	2.40
(standard errors)	(0.22)	(0.02)
adaptive kernel estimates of $\delta_a B(0)\mu/B$:	0.823	0.023

Turning to manual male workers (Table 15), we see that in each year the kernel estimate of I_1 is somewhat larger than that of $\delta_\alpha B(1)/B$. The OLS estimate of $\delta_\alpha B(1)/B$ is somewhat larger than the corresponding kernel estimate in all years except 1985. As was to be expected from the results for male manual labour supply (Table 9), the kernel estimates of $\delta_\alpha B(1)/B$ and I_1 are decreasing in the years from 1971 to 1977 and increasing from 1977 onwards. The values of $\delta_a B(0)\mu/B$ are as follows:

1971	1973	1975	1977	1979	1981	1983	1985
0.893	0.852	0.837	0.817	0.857	0.874	0.875	0.894

We have:

	arithmetic mean	standard deviation
adaptive kernel estimates of $\delta_\alpha B(1)/B$:	0.812	0.024
least-squares estimates of $\delta_\alpha B(1)/B$:	0.806	0.110
adaptive kernel estimates of I_1:	0.830	0.024
adaptive kernel estimates of $\delta_a B(0)$:	27.02	1.53
least-squares estimates of $\delta_a B(0)$:	25.60	4.00
(standard errors)	(0.36)	(0.01)
adaptive kernel estimates of $\delta_a B(0)\mu/B$:	0.862	0.025

Disregarding the exceptionally low value of $\delta_\alpha B(1)/B$ in 1985, we obtain for the OLS estimates the arithmetic mean 0.847 and the standard deviation 0.026.

Finally, Table 16 contains the estimates for the subgroup "non-manual male workers". As in the case of labour supply, the estimations were carried out separately on the eight annual samples. The kernel estimate of $\delta_\alpha B(1)/B$ is in each year slightly larger than that of I_1. The estimates are decreasing (resp. increasing) over the years 1971-1979 (resp. 1979-85). The first five OLS estimates of $\delta_\alpha B(1)/B$ are larger than the corresponding kernel estimates while the last three OLS estimates are smaller. Contrary to the kernel method, the linear regression model implies that the net earnings elasticity for non-manual male workers was strictly decreasing from 1971 to 1983. The OLS estimate of $\delta_\alpha B(1)/B$ for 1985 is larger (resp. smaller) than that for 1983 (resp. 1979). The kernel method, on the other hand, implies that $\delta_\alpha B(1)/B$ was slightly larger in 1985 than in 1979. The estimates of $\delta_a B(0)\mu/B$ are as follows:

1971	1973	1975	1977	1979	1981	1983	1985
0.837	0.832	0.790	0.756	0.765	0.777	0.795	0.789

We obtain the following means and standard deviations:

	arithmetic mean	standard deviation
adaptive kernel estimates of $\delta_\alpha B(1)/B$:	0.763	0.026
least-squares estimates of $\delta_\alpha B(1)/B$:	0.768	0.056
adaptive kernel estimates of I_1:	0.739	0.021
adaptive kernel estimates of $\delta_a B(0)$:	21.93	1.35
least-squares estimates of $\delta_a B(0)$:	21.86	2.63
(standard errors)	(0.35)	(0.04)
adaptive kernel estimates of $\delta_a B(0)\mu/B$:	0.793	0.027

Thus, the linear regression model fits the data for male workers considerably better than those for the total labour force. However, the OLS estimates fluctuate more than the kernel estimates. The kernel estimates of $\delta_\alpha B(1)/B$ for the group "male workers" deviate on average by around 2 per cent from its arithmetic mean; for manual and non-manual male workers the average deviation from the mean is approximately 3 per cent. Looking at the OLS estimates, we have deviations in the order of 6, 14 and 7 per cent, respectively. Hence, the OLS estimates are less reliable when considering single years.

An obvious drawback of the linear regression model is that it completely ignores distributional aspects of a wage increase. The kernel estimates, on the other hand, show that in all cases a uniform absolute increase in the wage rates leads to a somewhat larger rise in per capita net earnings than a uniform proportional increase.

It is a little puzzling that the results of the kernel estimations for full-time workers give us a ranking of the net earnings elasticities that does not agree with the ranking of the corresponding labour supply elasticities. The ranking of the net earnings elasticities obtained from the OLS estimates agrees very well with that of the labour supply elasticities. However, as was to be expected from the results for labour supply, both estimation methods lead to a small range of net earnings elasticities for full-time workers.

Table 11 Net Earnings

All workers and subsamples, 1983; adaptive kernel and OLS estimates

Standard errors of the OLS estimates in parentheses

							OLS-Estimates				
Sample	B	$\delta_a B(0)$	$\delta_a B(1)$	$\frac{\delta_a B(1)}{B}$	I_1	I_2	a	b	$b \cdot \bar{W}$	$\frac{b \cdot \bar{W}}{a + b \cdot \bar{W}}$	$\bar{W}$
All Workers	80204	25.92	75333	0.939	0.986	0.407E-3	15921	20.72	65972	0.806	3184
							(737)	(0.20)			
Females	51629	20.84	47628	0.923	0.952	0.471E-3	16327	14.41	35953	0.688	2495
							(872)	(0.30)			
Manual Females	36899	23.00	43591	1.181	1.149	0.648E-3	-5228	23.11	42730	1.139	1849
							(1527)	(0.79)			
Non-Manual Females	59839	18.59	47286	0.790	0.812	0.361E-3	24387	12.63	35996	0.596	2850
							(1200)	(0.36)			
Males	102961	23.26	80726	0.784	0.778	0.263E-3	30932	19.91	74085	0.705	3721
							(1010)	(0.24)			
Manual Males	88071	25.82	73406	0.833	0.844	0.333E-3	11452	26.04	77729	0.872	2985
							(1119)	(0.35)			
Non-Manual Males	123304	21.08	94178	0.764	0.739	0.194E-3	40728	18.13	84286	0.674	4649
							(1993)	(0.37)			
Manuals	70854	30.47	76223	1.076	1.136	0.521E-3	-6098	29.96	78136	1.085	2608
							(893)	(0.32)			
Non-Manuals	88741	22.83	78633	0.886	0.894	0.318E-3	19794	19.18	70793	0.781	3691
							(1171)	(0.27)			

Table 12 All Workers: Net Earnings
Adaptive kernel and OLS estimates[1]; global window width:
h = 0.12·(mean gross wage rate)

							OLS-Estimates				
YEAR	B	$\delta_a B(0)$	$\delta_a B(1)$	$\frac{\delta_a B(1)}{B}$	I_1	I_2	a	b	$b \cdot \bar{W}$	$\frac{b \cdot \bar{W}}{a + b \cdot \bar{W}}$	$\bar{W}$
1970	16130	30.22	14918	0.925	0.968	0.235E-2	8846	14.13	7701	0.465	545
							(168)	(0.24)			
1971	17483	29.77	16223	0.928	0.974	0.215E-2	1721	27.15	16263	0.904	599
							(117)	(0.17)			
1972	19923	29.56	18278	0.917	0.957	0.188E-2	2027	27.27	18298	0.900	671
							(140)	(0.18)			
1973	22317	29.12	20518	0.919	0.958	0.161E-2	3216	25.57	19817	0.860	775
							(185)	(0.20)			
1974	25671	29.00	24364	0.949	0.985	0.138E-2	6030	22.34	20352	0.771	911
							(218)	(0.20)			
1975	31312	27.33	29413	0.939	0.974	0.106E-2	3378	24.64	28582	0.894	1160
							(239)	(0.18)			
1976	35244	26.60	33075	0.938	0.979	0.902E-3	8025	20.55	27743	0.776	1350
							(303)	(0.20)			
1977	39308	27.13	37325	0.950	0.992	0.823E-3	6124	22.72	33876	0.847	1491
							(330)	(0.20)			
1978	45595	27.46	42484	0.932	0.979	0.725E-3	11259	20.69	35173	0.758	1700
							(386)	(0.20)			
1979	52251	28.20	49842	0.954	0.996	0.651E-3	8647	23.38	44679	0.838	1911
							(449)	(0.21)			
1980	62666	27.86	59872	0.955	0.997	0.543E-3	7954	24.16	56220	0.876	2327
							(497)	(0.19)			
1981	70188	26.80	66669	0.950	0.993	0.475E-3	24631	17.37	47177	0.657	2716
							(617)	(0.19)			
1982	75021	26.41	70265	0.937	0.996	0.448E-3	9593	22.82	66520	0.874	2915
							(601)	(0.18)			
1983	80204	25.92	75333	0.936	0.986	0.407E-3	15921	20.72	65972	0.806	3184
							(737)	(0.20)			
1984	83710	25.53	77156	0.922	0.971	0.380E-3	25687	17.90	59840	0.700	3343
							(772)	(0.19)			
1985	92198	25.45	85433	0.927	0.958	0.345E-3	29186	17.84	65437	0.692	3668
							(901)	(0.20)			

1) Standard errors of the OLS estimates in parentheses.

Table 13 Full-Time Workers, 1983: Net Earnings
Adaptive kernel and OLS estimates[1)]

Sample	B	$\delta_a B(0)$	$\delta_e B(1)$	$\frac{\delta_e B(1)}{B}$	I_1	I_2	OLS-Estimates a	b	$b\cdot\bar{W}$	$\frac{b\cdot\bar{W}}{a+b\cdot\bar{W}}$	$\bar{W}$
All Workers (Full-Time)	94858	22.97	76269	0.804	0.779	0.268E-3	15273 (648)	24.00 (0.17)	81576	0.842	3399
Females	68939	19.32	50693	0.735	0.704	0.298E-3	12253 (645)	22.02 (0.23)	57979	0.826	2633
Manual Females	57894	19.90	40773	0.704	0.680	0.357E-3	12406 (1117)	22.55 (0.51)	46115	0.788	2045
Non-Manual Females	72901	19.18	54504	0.748	0.715	0.278E-3	11420 (838)	22.16 (0.27)	62912	0.846	2839
Males	106016	22.46	81324	0.767	0.744	0.233E-3	21348 (880)	23.31 (0.21)	86830	0.803	3725
Manual Males	91313	22.45	67125	0.735	0.720	0.259E-3	15696 (1104)	25.23 (0.34)	76724	0.830	3041
Non-Manual Males	125805	21.28	96783	0.769	0.736	0.184E-3	23137 (1722)	22.86 (0.33)	105087	0.820	4597
Manuals	85703	23.59	66545	0.776	0.762	0.294E-3	12014 (903)	26.01 (0.29)	74831	0.862	2877
Non-Manuals	103393	22.60	86811	0.840	0.803	0.245E-3	12985 (1013)	24.01 (0.23)	92751	0.817	3863

Table 14 Male Workers: Net Earnings
Adaptive kernel and OLS estimates[1)]; global window width:
h = 0.12·(mean gross wage rate)

YEAR	B	$\delta_a B(0)$	$\delta_e B(1)$	$\frac{\delta_e B(1)}{B}$	I_1	I_2	OLS-Estimates a	b	$b\cdot\bar{W}$	$\frac{b\cdot\bar{W}}{a+b\cdot\bar{W}}$	$\bar{W}$
1971	22290	26.98	17507	0.785	0.796	0.143E-2	4124 (149)	26.39 (0.18)	18711	0.819	709
1973	28872	26.17	22244	0.770	0.771	0.104E-2	6939 (262)	24.63 (0.24)	22733	0.766	923
1975	40045	23.82	30103	0.752	0.753	0.670E-3	9669 (313)	22.93 (0.21)	31024	0.762	1353
1977	49950	22.73	36473	0.730	0.730	0.506E-3	11924 (403)	22.66 (0.21)	38817	0.765	1713
1979	66810	24.36	50421	0.755	0.757	0.410E-3	15101 (548)	23.80 (0.22)	52836	0.778	2220
1981	90432	23.54	69743	0.771	0.767	0.299E-3	25605 (754)	20.96 (0.21)	66464	0.722	3171
1983	102961	23.26	80726	0.784	0.778	0.263E-3	30932 (1010)	19.91 (0.24)	74085	0.705	3721
1985	119581	22.87	92500	0.774	0.760	0.223E-3	41890 (1294)	18.55 (0.25)	80117	0.657	4319

Table 15 Male Manual Workers: Net Earnings
Adaptive kernel and OLS estimates[1]; global window width:
h = 0.09·(mean gross wage rate)

YEAR	B	$\delta_a B(0)$	$\delta_a B(1)$	$\frac{\delta_a B(1)}{B}$	I_1	I_2	OLS-Estimates a	OLS-Estimates b	$b\cdot\bar{W}$	$\frac{b\cdot\bar{W}}{a+b\cdot\bar{W}}$	$\bar{W}$
1971	20132	30.07	16722	0.831	0.866	0.176E-2	2560	29.74	17785	0.874	598
							(211)	(0.33)			
1973	26261	28.81	21042	0.801	0.824	0.124E-2	4491	28.30	21989	0.830	777
							(295)	(0.36)			
1975	37105	26.26	29235	0.788	0.807	0.798E-3	7315	25.43	30084	0.804	1183
							(449)	(0.36)			
1977	45757	25.17	35008	0.765	0.786	0.612E-3	8276	25.54	37927	0.821	1485
							(564)	(0.36)			
1979	60657	27.14	49199	0.811	0.826	0.503E-3	8516	27.61	52873	0.861	1915
							(754)	(0.37)			
1981	79190	26.35	65595	0.828	0.841	0.379E-3	10753	26.46	69510	0.866	2627
							(1006)	(0.36)			
1983	88071	25.82	73406	0.833	0.844	0.333E-3	11452	26.04	77729	0.872	2985
							(1119)	(0.35)			
1985	102987	26.51	86141	0.836	0.849	0.297E-3	49980	15.68	54441	0.521	3472
							(1505)	(0.38)			

Table 16 Male Non-Manual Workers: Net Earnings
Adaptive kernel and OLS estimates[1]; global window width:
h = 0.14·(mean gross wage rate)

YEAR	B	$\delta_a B(0)$	$\delta_a B(1)$	$\frac{\delta_a B(1)}{B}$	I_1	I_2	OLS-Estimates a	OLS-Estimates b	$b\cdot\bar{W}$	$\frac{b\cdot\bar{W}}{a+b\cdot\bar{W}}$	$\bar{W}$
1971	27240	24.32	21793	0.800	0.779	0.103E-2	3273	26.34	24707	0.833	938
							(319)	(0.29)			
1973	34615	23.98	27371	0.791	0.761	0.790E-3	6470	24.38	29280	0.819	1201
							(622)	(0.42)			
1975	45925	21.85	35008	0.762	0.741	0.529E-3	9107	22.61	37555	0.805	1661
							(586)	(0.32)			
1977	58338	20.68	42479	0.728	0.712	0.389E-3	11869	22.15	47246	0.799	2133
							(799)	(0.34)			
1979	77728	21.77	56499	0.727	0.716	0.312E-3	16830	22.72	62071	0.787	2732
							(1045)	(0.34)			
1981	108087	21.06	80563	0.745	0.721	0.218E-3	33646	19.11	76249	0.694	3990
							(1452)	(0.32)			
1983	123304	21.08	94178	0.764	0.739	0.194E-3	40728	18.13	84286	0.674	4649
							(1993)	(0.37)			
1985	144111	20.71	112979	0.784	0.741	0.160E-3	39668	19.42	106596	0.729	5489
							(2568)	(0.41)			

1) Standard errors of the OLS estimates in parentheses.

7. Relation to the Literature

We mentioned already in the Introduction that there is a large body of empirical work on labour supply. We want to give only a very brief overview here. A comprehensive discussion of the literature is provided by Killingsworth (1983); the interested reader will find there more than 500 references, covering some 20 years of research on labour supply.

Contrary to our approach, the literature has been concerned with estimating the parameters of fully specified functional forms. The labour supply function has been estimated either directly or indirectly via the estimation of a commodity demand system. In the first case the studies have chosen a functional form and then computed the regression of labour supply on the wage rate (w), non-labour income (m) and a number of other observable personal characteristics which we will not consider here.
Functional forms which have been used in the literature include:

(1) $l(w,m) = \alpha_1 + \alpha_2 \cdot w + \alpha_3 \cdot w^2 + \alpha_4 \cdot m + \alpha_5 \cdot m^2 + \alpha_6 \cdot w \cdot m$ (quadratic labour supply function),

(2) $l(w,m) = \alpha_1 + \alpha_2 \cdot \log w + \alpha_3 \cdot m$ (semi-log labour supply function),

(3) $l(w,m) = \alpha_1 + \alpha_2 \cdot (m/w) + \alpha_3/w$ (linear earnings function),

(4) $wl(w,m)/m = \alpha_1 + \alpha_2 \cdot \log w + \alpha_3 \cdot \log m$ (share linear in logarithm),

where the α_i are the parameters to be estimated.

In the second case, where $l(w,m)$ is not estimated directly, one specifies the functional form of the commodity demand function $(p,w,m) \longmapsto f(p,w,m) \in R^n_+$ (p and n denote, respectively, the commodity price system and the number of commodities). The parameters of the demand function are usually estimated by the method of maximum likelihood; the labour supply function is then obtained from the budget identity $wl(p,w,m)+m = pf(p,w,m)$. In virtually all empirical work the demand function is modelled in the following additive form:

$$f(p,w,m) = \sum_{j=1}^{s} g_j(p,w) \cdot v_j(m),$$

where the v_j are real-valued functions and $g_j(p,w) \varepsilon R^n$ $(j=1,\ldots,m)$. Special cases of this form are polynomial expenditure systems [i.e., $p_i f_i$ is a polynomial in (w,m)] and the well-known "almost ideal demand system" (see Deaton and Muellbauer, 1980). We remark that the labour supply function (4) is essentially the AIDS form for the case of labour supply.

In particular, the literature has focused on a question which we have not pursued here: Can the estimated commodity demand and labour supply system be looked at as the commodity demand and labour supply function of a fictitious "representative" consumer? This question can be answered by computing the matrix of substitution effects and checking whether the estimated matrix is symmetric and negative semi-definite (the standard reference is Hurwicz and Uzawa, 1971); if one considers only two goods, i.e, labour and total income = wl(w,m) + m, then the labour supply function is consistent with the utility maximisation hypothesis if and only if $\delta_w l(w,m) - l(w,m)\delta_m l(w,m) \geq 0$ for all w and m. However, for the purpose of the present study the notion of a "representative consumer" is not helpful; furthermore, the market demand function may have properties which an individual demand function does not have (see Chapter 1). We therefore agree with Hildenbrand (1983) that this concept "might be misleading".

In Section 2 we have explained why we have not regressed labour supply on the marginal net wage rate. However, if one constructs the labour supply function via the estimation of the corresponding commodity demand system, then "it is necessary to adjust the pre-tax hourly wage rate for the presence of taxes on earnings...to preserve the aggregate budget identity relating expenditures to income" (Abott and Ashenfelter 1976, p. 395); and this is usually done in the literature.

The functional forms (1)-(4) and other specifications of the labour supply function have been investigated by Stern (1986); the author discusses, among other things, consistency with utility maximisation and the "flexibility" of the functions (i.e., their response to changes in the wage rate). Figure 11 on the next page shows two polynomial least-squares estimates and an adaptive kernel estimate of the aggregate labour supply function for 1983. The polynomials are of degree n=4, 7, respectively; the adaptive kernel smoother was computed with h=380. As we see, the empirical regression curve cannot be very well approximated by a polynomial of low order.

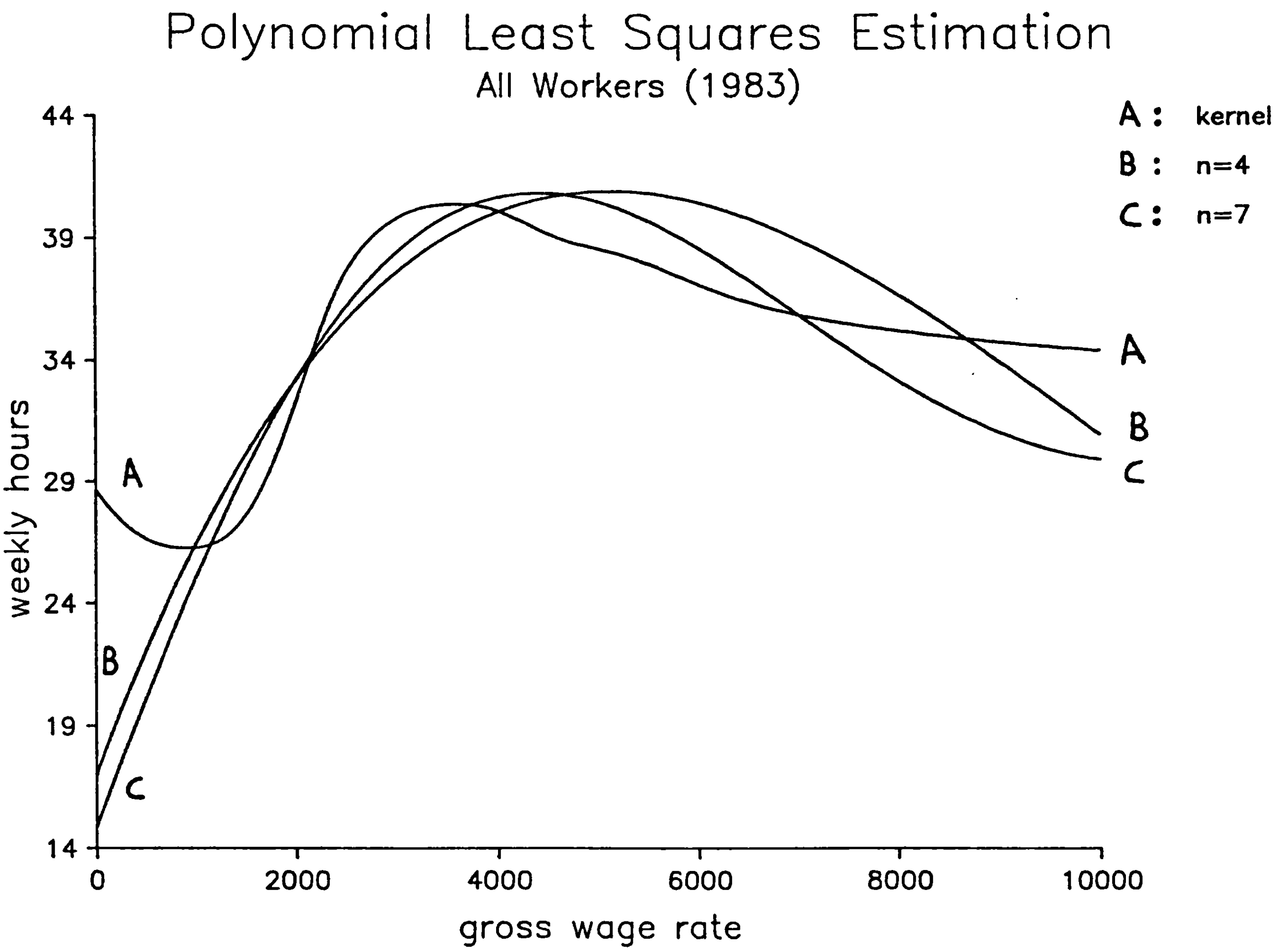


Figure 11

Let $\hat{l}(w,m)$ denote the estimated functional form. The elasticity of labour supply with respect to the wage rate is computed in the literature by substituting the sample means of w and m in $\delta_w\hat{l}(w,m)[w/\hat{l}(w,m)]$. Notice that this value may differ considerably from the average elasticity

$$\iint \delta_w\hat{l}(w,m)\,[w/\hat{l}(w,m)]\cdot\rho(w,m)\,dw\,dm,$$

where $\rho(w,m)$ denotes the joint density of w and m.

It is interesting to observe that our estimates have confirmed the famous "backward bending" labour supply function. Many studies on male labour supply have computed elasticities that are similar to our results (a useful survey is Pencavel, 1986). Let us mention four cross-section studies. The first three studies focus on the joint determination of labour supply and commodity demands; the fourth study does not consider several consumption goods but only the aggregate "total household expenditure" (recall that our estimates of $\delta_\alpha L(1)/L$ for males are ranging from -0.098 to +0.049):

1) Atkinson and Stern (1981) model that consumption activity involves time. For the empirical implementation of the model they use the so-called *Stone-Geary* utility function. In the standard neoclassical model of a household, where the time aspect of consumption is not explicitly treated, this utility function leads to a linear expenditure system (LES), i.e., the commodity demand function f(p,w,m) is linear in (w,m); by the budget identity this, in turn, implies the labour supply function (3). However, if consumption takes time, then the labour supply behaviour of an individual with Stone-Geary utility function may be quite complex (Figure 9 in Stern, 1986, illustrates this).
The sample is taken from the 1973 FES and consists of 1617 households with a male head aged 18-64 who is in full-time employment and whose gross hourly wage is in the range between £0.85 and £3.0. Nine commodity categories are considered. The authors first estimated expenditure equations of the form $p_i f_i = \alpha_i + \beta_i w + \mu_i m$ (i=1,...,9) by the method of least-squares and obtained a labour supply elasticity of -0.146 (evaluated at the sample means of the explanatory variables); Atkinson and Stern remark that "this value is within the range of estimates...typically found in empirical studies" (p. 288). The extended LES (where consumption takes time) was then

estimated by the method of maximum likelihood; in the extended model the elasticity is -0.230.

The next two studies consider both male and female labour supply:

2) By generalising the simple LES resulting from the *Stone-Geary* utility function, Blundell and Walker (1982) obtain a model for the consumption and labour supply decisions of two persons, i.e., husband and wife. Their data are taken from the 1974 FES. The sample consists of 103 households which were chosen so that the head of each household is a male manual worker whose wife is in paid employment. The authors compute an elasticity of male labour supply of -0.286 (the commodity demand system consists of 6 commodity aggregates).

3) Using data for the Federal Republic of Germany for the year 1984 and estimating essentially a linear expenditure system (with nine commodity groups), Kaiser (1990) obtains for males an elasticity of -0.014 (the sample consists of 2223 households; each man works more than 10 hours per week).

The above authors have estimated static labour supply models as we have done in this chapter. One may argue that this is not the appropriate model for analysing labour supply decisions. At the end of their article, Blundell and Walker (1982) point out that "our estimates may be picking up lifecycle phenomena" (p. 363). In a subsequent work they returned to this question:

4) Blundell and Walker (1986) estimate a life-cycle model of family labour supply; the data are extracted from the 1980 FES. The sample consists of 1378 households which were chosen so that (i) each houshold consists of two married working employees, and (ii) the head of the household is either a manual worker, a shop assistant or a clerical worker. The estimated model produces an elasticity of male labour supply of -0.263.

The elasticities computed by Atkinson and Stern (1981) and Blundell and Walker (1982, 1986) differ considerably from our estimates. Other studies have obtained elasticities which are less far away from our results. Generally speaking, there is agreement in the literature that the labour supply elasticity for males is "negative and small", but there is not much agreement on the exact magnitude of the elasticity. A comparison of the above results with a well-known time-series study may be interesting:

6) Abbott and Ashenfelter (1976, 1979) estimate the demand functions for seven commodity aggregates and the corresponding labour supply function using four functional specifications and aggregate time-series data for the United States; the data cover the years 1929-67 and are expressed in per capita terms. The authors obtain the following labour supply elasticities: -0.143 (Rotterdam model), -0.070 (separable Rotterdam model), -0.070 (Stone-Geary utility function; i.e., LES) and +0.879 ("addilog" indirect utility function); the figures are taken from their correction (1979).

It is usually argued in the literature that the labour supply function for females is positive sloped and that female labour supply responds much more to changes in the wage rate than male labour supply. As in the case of male labour supply there is, however, not much agreement on the magnitude of the female labour supply elasticity: "the range of estimates for the uncompensated own-wage elasticity for women is, if anything, larger than the range of estimates for men: between 0.200 and 0.900 in most aggregate cross-section and microlevel cross-section studies of female labor supply" (Killingsworth, 1983, Chapter 3, pp. 103-104). For instance, Blundell and Walker (1982) compute the following elasticities for married women: 0.427 (no children), 0.107 (one child; age = 3) and -0.193 (2 children; age = 3 and 6); and in Kaiser's (1990) study the labour supply elasticity of married women is 0.52. (An interesting volume of readings on female labour supply is Smith, 1980. See also Killingsworth and Heckman, 1986.)

In Section 6 of this study we have also seen that the elasticity of labour supply is substantially higher in the population of "all females" than in that of "all males" (Table 4); the aggregate labour supply curve for females estimated in Section 4 is increasing on the interval [$w_{0.05}$, $w_{0.75}$] (Figure 6, Table 1a). However, our results for full-time workers do not provide support for the view that "male labor supply is much less sensitive to wage changes than is female labor supply" (Killingsworth, 1983, Chapter 3, p. 102): the labour supply curve for the group "full-time female workers" is decreasing on $[0, w_{0.99}]$ with an elasticity close to zero (Figure 7, Table 7).

In this work the regression curves were estimated on samples of workers, and the wage rate was constructed by dividing gross earnings by hours of work. We mentioned in the Introduction that a sample of workers is in general not a random sample from the total population of individuals whose labour supply function one wants to estimate. More precisely, one considers only individuals having a market wage that exceeds their reservation wage. Two further problems are: "endogeneity of the explanatory varable" and "spurious correlation between the wage rate and hours due to measurement errors".

These issues have received much attention in the literature. It has been suggested that one should replace the observed market wage by an ***instrumental variable.*** Such an instrumental variable can be obtained by regressing the observed wage rate on a number of personal characteristics (such as age, education and work experience). Let $(x_1,...,x_m)$ be a vector of personal attributes, and let $\alpha_0,...,\alpha_m$ be the estimated coefficients obtained from the regression $w = a_0+a_1x_1+...+a_mx_m$, where w = earnings/hours. An instrument for w is then given by $w_I = \alpha_0+\alpha_1x_1+...+\alpha_mx_m$. By using this "wage function", one can also compute "imputed" wage rates for those individuals in a given sample who are not in paid employment and hence has data for the entire population (see Killingsworth, 1983, Chapter 3, p. 92, for references).

It would be interesting to explore whether the shape of the labour supply curves estimated in this chapter is robust with respect to alternative definitions of the wage variable. Unfortunately, the FES lacks information on work experience and education.

Our empirical study of the labour market was motivated by the work of K. Hildenbrand and W. Hildenbrand (1986) who analysed the matrix of mean income effects (see Chapter 1) using nonparametric estimates for the Engel curves and the income density. Härdle and Jerison (1988) estimated Engel curves on the FES sample of "all households" for each odd numbered year from 1969 to 1977. Their observations are similar to those which we made in Section 5: "Real Engel curves (with quantity demanded and real total expenditure on the axes) vary over time, but their shapes are generally quite stable. Mean normalized Engel curves...are found not to vary greatly over time" (the quotation is taken from their Abstract).

8. Concluding Remarks

Using nonparametric smoothing techniques which have not yet received much attention in the literature (the data analysis was carried out in 1990), we have tried to answer an old question: Does a wage increase lead to an increase in aggregate labour supply? Our estimates of the labour supply elasticity confirm the results of many previous studies and may therefore be reassuring to economists who prefer parametric modelling. However, if we exclude part-time workers then our findings for females are at variance with the literature; in the case of male workers there are also studies which have produced estimates that differ considerably from our results (see Killingsworth, 1983, Chapters 3 and 4).

It is, of course, difficult to compare studies which have used different data sets and sample selection rules. We have not explored what happens if one applies the methods usually used in the literature to the data from which we have constructed our estimates. In some cases this may have quite an effect on the estimated labour supply elasticity. At the very least, we think, results based on a particular functional form should be interpreted with care.

In Section 2 we have assumed that external effects on the labour market can be disregarded, i.e., the labour supply of an individual $i \varepsilon I$ at the wage rate w(i) does not depend on the labour supply and the wage rate of any other individual i' in I. This is, of course, a standard assumption.

On the other hand, sociological explanations of the functioning of the labour market suggest that the labour supply of individuals is essentially determined by their conception of a "fair wage"; whether or not individual i thinks that w(i) is a "fair wage", in turn, depends on the distribution of wages in his or her "work group" (see Akerlof, 1982, for details and a simple model along these lines). It is natural to assume that the wage rates earned by other people play a role in explaining the labour supply of a person. In other words, the aggregate labour supply function $l(w)$ may depend on the wage rate density ρ, i.e., $l(w)=l(w,\rho)$.

This means that a small variation in the wage rates may have two effects on aggregate labour supply, namely a "direct" effect via the change in w (which we have explored here) and an "indirect" effect via the change in ρ. Let us be more specific. Suppose $l(w)$ depends on the *inequality* in the wage rate distribution and that the inequality in the wage rates is represented by the ratio of standard deviation to mean. Then the inequality decreases if we add to all wage rates an amount $a>0$. In the case of a proportional wage increase, however, the inequality does not change. Thus, an absolute wage increase will have two effects on labour supply while a proportional increase will have only a direct wage effect.

In our opinion, the "fair wage" argument is indeed important when looking at individual labour supply decisions within firms. However, in this study we were concerned only with decisions about hours of work (it is open to discussion whether this is a relevant measure of labour supply); within a firm one should consider the number and the quality of the tasks performed by an employee.

Since our approach to the FES data is the same as that of K. Hildenbrand and W. Hildenbrand (1986), who estimate Engel curves, we can compare their results with our findings. It appears that the response of individuals to a change in the wage rate differs substantially from their response to a change in the consumer goods prices. The authors write: "We computed the kernel estimator for various kernel functions...and a large number of commodities in the sample of the Family Expenditure Survey for the years 1969 to 1981. For certain commodities, like fuel, alcohol and butter, and in particular for all aggregates, like food, housing, clothes and footwear, and services, these kernel estimators supported well the hypothesis that the Engel curve is monotone increasing" (p. 258).

This has the following implication: If the average value of the individual substitution effects has a negative sign (we do not have to assume that all individuals behave "rationally") and if the dispersion of demand around the estimated Engel curve is an increasing function of household income, then an increase in the relative price of an aggregate commodity leads to a decrease in demand (see W. Hildenbrand, 1989a, for details; especially Proposition 2, p. 271). On the other hand, the labour supply curves for full-time workers are downward sloping, i.e., it seems that full-time workers do not reduce their consumption of leisure if leisure becomes more expensive.

However, the elasticity of full-time labour supply with respect to the wage rate is close to zero. Disregarding distributional aspects, our estimations therefore suggest that full-time labour supply is an interesting object for taxation. The qualitative picture changes if on considers the total population of workers (i.e., full-time plus part-time workers). Our estimations suggest an elasticity of total labour supply of around 0.2. In other words, it appears that in the entire population of workers the "law of supply" is valid.

In Chapter 0 we have seen that the demand for labour of a profit maximising firm is a decreasing function of the wage rate. We do not know whether the elasticity of aggregate demand for full-time labour is much higher than that of full-time labour supply. If this is the case, then a small proportional increase in all wage rates will imply an increase in excess labour supply (= labour supply minus labour demand), i.e., the excess supply for full-time labour will satisfy the "law of supply". (One should not forget that the framework of Chapter 0 is very simple. If an increase in the wage rate implies an increase in the quality of labour supplied by individuals, then the profit maximising firm may decide to demand more labour. However, there are no "efficiency wages" in Chapter 0.)

We would hesitate to draw policy conclusions from our findings. An extremely simple labour supply model was estimated in this chapter; in addition, our data analysis has shortcomings. We mentioned already in the Introduction three problems, namely: endogeneity of the explanatory variable, errors in the variables and sample selection bias. Furthermore, it is unsatisfactory that the regression curves were estimated on very heterogeneous data sets: "those who are observed, say, to work more and receive lower wages may differ in some systematic way from those who work less and receive higher wages. What we have to do is to specify the characteristics in which people differ, and this is rarely considered in detail" (Atkinson and Stiglitz, 1980, p. 52). Finally, we have said nothing about the standard errors of the estimates. It would be interesting to test micro-econometric labour supply models by constructing confidence bands around a kernel smoother (details can be found, for example, in Härdle, 1990, Chapter 4, pp. 98-109). Hence, there are several directions for future research in this area.

9. Notes

1) Lucas and Rapping (1970) take the view that unemployment is essentially a form of leisure time, i.e, there is no unvoluntary unemployment.

2) "The sets which constitute the universe of discourse must always be explicitly listed at the outset" (Debreu, 1959, p. 3). In order to avoid going into technicalities we have not done this here. To define a distribution on the space $I \times R_+$, one needs a measurable structure on I. If I is the space of preference relations, then I can be endowed with a metric; i.e., μ is a normed Borel measure on the product space $I \times R_+$. Mathematically, labour supply is represented by a function $l: I \times R^n_+ \times R_+ \times T \to R_+$, where T denotes the set of tax functions. Hence, $l^i(p,w,t) = l(i,p,w,t)$. For given (p,t), the function $l(\cdot,p,\cdot,t): I \times R_+ \to R_+$ is assumed to be μ-integrable, that is, $\int l(\cdot,p,\cdot,t)d\mu < \infty$; see Hildenbrand (1974) for mathematical details (especially p. 96).

3) We do not assume that the individual labour supply functions are differentiable with respect to w; the functions $l^i(p,\cdot,t): R_+ \to R_+$ may have jumps and kinks. Strictly speaking, we merely assume that averaging over a large set of individuals produces a smooth curve. The idea that demand and supply functions may be smoothed by aggregation is very old (see, e.g., Cournot, 1838, pp. 49-50). The mathematical conditions under which this is possible have been studied in the literature; the standard reference is Trockel (1983).

4) Let us give a simple example: Suppose an individual wants to maximise his net labour income subject to a cost function $K(l)$. (Usually $K(l)$ is interpreted as the "disutility of effort".) The first-order condition for a maximum of $N(l) = wl - t(wl) - K(l)$ is $\{1-t'(wl)\}w = K'(l)$. Let us assume that the optimisation problem has a unique solution $l^* = l^*(w,t)$. Setting $w' = \{1-t'(wl^*)\}w$, then l^* will also maximise the function $w'l - K(l)$. Hence, denoting by $l^\circ = l^\circ(w')$ the maximum for the latter function, we have $l^*(w,t) = l^\circ(w')$ (note that this is the message of Figure 1 in Section 2).

5) Alternatively, one could estimate the unknown regression function by the penalised likelihood method discussed in Chapter 2. As in the case of density estimation, a natural roughness penalty is $R(g) = \int g''(w)^2 dw$. If the residuals $y_i - g(w_i)$ $(i=1,\dots,n)$ are independent, normally distributed with zero mean and common variance σ^2, then their log-likelihood is given by $(-1/2\sigma^2)\Sigma[y_i - g(w_i)]^2 - (n/2)\log(2\pi\sigma^2)$. It is interesting to remark that, for given $\alpha > 0$, the "penalised" sum of squares $\Sigma[y_i - g(w_i)]^2 + \alpha \cdot R(g)$ will be minimised by a cubic spline (see, e.g., Eubank, 1988, Chapter 5; or Silverman, 1985).

6) The following interpretation of $\Sigma y_i \frac{1}{h} K(\frac{w-w_i}{h})$ may be helpful. Suppose that there are n workers in the economy. For any given wage rate w, the decision problem of individual i is either to supply y_i units of labour or to stay away from the labour market. Let $P_i(w)$ be the probability that individual i supplies y_i when faced with w. Expected labour supply at the wage rate w is then given by $\Sigma y_i P_i(w)$. Suppose K is a "bell-shaped" kernel and $P_i(w) = (1/h)K((w-w_i)/h)$. This has then the following implications for the probability of working: (i) $P_i(w_i) > P_i(w)$ for all $w \neq w_i$; (ii) $P_i(w_i)$ decreases as h becomes larger; (iii) if w differs "very much" from w_i, then $P_i(w)$ increases as h becomes larger.

7) Eubank (1988, p. 112) remarks: "...the use of kernels with support on the entire line results in estimators with global bias difficulties which are localized to the boundaries when K has finite support. Thus kernels with finite support seem preferable". In this work we set K(w) equal to zero for all w such that $|w|>5$. Technically speaking, this means that we used the on [-5,5] truncated Gaussian kernel.

8) The optimal window width at a given point w depends on the value of the unknown probability density ρ at w and the local behaviour of the unknown regression function $g(\cdot)$ around w. If $g(\cdot)$ is locally linear at w [i.e., $g''(w)=0$], then all observations near w will provide information about the behaviour of $g(\cdot)$. Conversely, if $g(\cdot)$ is changing rapidly near w [i.e., $|g''(w)|$ is large], then only data points very close to w will do so. This suggests that the globally optimal window width will be a decreasing function of the "average value" of $|g''(w)|$. Note that exactly this is the message of the formula for the optimal window width (from the point of view of minimising the integrated mean squared error) of the density estimator given on page 76 : h^* is a strictly decreasing function of $\int \rho''(x)^2 dx$. In the case of regression estimation one obtains the same relation between kernel function, sample size and local variability of the unknown curve, on the one hand, and the (globally) optimal window width on the other (see, e.g., Eubank, 1988, Theorem 4.2, p. 135; and pp. 151-153).

9) By definition of the kernel smoothers $\hat{g}(w)$ (= regression) and $\hat{\rho}(w)$ (= density),

$$L = \int_{-\infty}^{+\infty} \hat{g}(w)\hat{\rho}(w)dw = \text{sample mean of the dependent variable.}$$

However, we integrate only over $[0,w_{0.99}]$. The kernel estimates of L therefore deviate very slightly from the sample means given in the Appendix.

Appendix

The Evolution of Some Sample Statistics over the Period 1970-1985

The Appendix presents some summary statistics for the earnings and labour supply data in the annual samples of "all workers" (Table 1) and in eight subsamples: female workers, female manual (resp. non-manual) workers, male workers, male manual (resp. non-manual) workers, manual workers and non-manual workers (Tables 2-9).

The summary measures considered here are: median, mean, standard deviation, coefficient of variation and skewness; the correlation between labour supply (l) and gross wage rate (w), and the correlation between net earnings (b) and gross wage rate.
The following notation is used (for a discussion of the FES data see Section 2 of Chapter 2; in particular pages 68-69):

GRWAGE	normal hourly gross wage (i.e., the gross wage rate)
HOURS	normal weekly hours of work including paid overtime
NEARN	normal net weekly earnings
MEAN	arithmetic mean of the observations
STD	sample standard deviation
CV	coefficient of variation (CV = STD/MEAN)
COR(l,w)	empirical correlation coefficient between HOURS and GRWAGE
COR(b,w)	empirical correlation coefficient between NEARN and GRWAGE
SKEWNESS	skewness of the empirical distribution, i.e., the third central moment of the data divided by the third power of the sample standard deviation

The calculations were carried out by SAS. The FES earnings data were originally recorded in tenths of pence per week. In the tables MEAN, MEDIAN and STD were converted into £ with two decimal places. When dividing the sample standard deviation by the sample mean, one will therefore obtain a value which deviates very slightly from the coefficient of variation calculated by SAS.

Table 1a All Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.48	0.54	0.45	83.19	40	38.29	12.52	32.71	0.082
1971	0.52	0.60	0.37	62.30	40	37.62	12.05	32.04	0.168
1972	0.59	0.67	0.39	58.78	40	37.60	12.27	32.62	0.180
1973	0.68	0.78	0.52	67.17	40	37.20	12.49	33.56	0.133
1974	0.81	0.91	0.58	63.65	40	37.09	12.76	34.39	0.170
1975	1.04	1.16	0.62	53.53	40	36.99	12.62	34.10	0.195
1976	1.21	1.35	0.75	55.20	40	36.55	12.67	34.68	0.139
1977	1.34	1.49	0.77	51.60	40	36.66	12.76	34.81	0.158
1978	1.51	1.70	0.95	56.12	40	36.61	12.56	34.32	0.121
1979	1.71	1.91	1.03	53.92	40	36.81	13.02	35.38	0.175
1980	2.05	2.33	1.32	56.92	39	36.27	12.93	35.65	0.173
1981	2.36	2.72	1.82	67.15	38	35.82	12.64	35.29	0.126
1982	2.53	2.92	1.67	57.35	38	35.54	12.83	36.09	0.174
1983	2.75	3.18	1.98	62.17	38	35.57	13.08	36.76	0.135
1984	2.88	3.34	2.16	64.58	38	35.44	13.16	37.13	0.134
1985	3.09	3.67	2.55	69.43	38	35.78	13.42	37.51	0.114

Table 1b All Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	SKEWNESS GRWAGE	SKEWNESS HOURS
1970	5370	15.32	16.54	10.16	61.43	0.630	1.48	24.23	-0.59
1971	8307	16.75	17.99	11.60	64.45	0.874	2.93	3.68	-0.70
1972	8168	19.00	20.31	12.51	61.59	0.859	1.93	2.72	-0.55
1973	8244	21.50	23.03	16.28	70.70	0.817	9.29	6.16	-0.61
1974	7621	24.53	26.38	16.50	62.55	0.785	2.70	9.88	-0.42
1975	8334	30.12	31.97	18.44	57.69	0.830	1.45	2.56	-0.41
1976	8137	34.00	35.77	20.21	56.51	0.758	1.23	4.48	-0.51
1977	8184	38.30	39.99	22.19	55.48	0.788	1.22	2.74	-0.45
1978	7805	44.37	46.44	25.86	55.68	0.764	1.29	4.11	-0.39
1979	7638	50.50	53.32	30.44	57.10	0.791	1.32	3.54	-0.37
1980	7846	60.00	64.18	38.71	60.31	0.827	1.57	3.63	-0.30
1981	8103	66.87	71.80	44.28	61.67	0.715	1.76	11.62	-0.39
1982	7774	71.04	76.12	46.38	60.94	0.823	1.76	3.14	-0.48
1983	6833	75.85	81.90	52.12	63.64	0.787	1.87	3.85	-0.34
1984	7089	80.00	85.53	52.30	61.15	0.739	1.51	6.82	-0.34
1985	6947	87.46	94.62	62.44	65.99	0.728	2.67	7.19	-0.25

Table 2a Female Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.32	0.38	0.23	59.99	35	29.98	12.13	40.45	-0.007
1971	0.36	0.42	0.25	58.86	35	29.80	11.96	40.11	0.018
1972	0.42	0.48	0.26	54.43	35	29.49	12.15	41.20	0.028
1973	0.48	0.55	0.32	57.06	34	28.89	12.17	42.13	-0.017
1974	0.60	0.68	0.53	77.30	32	28.85	11.88	41.16	0.022
1975	0.79	0.88	0.46	52.02	33	28.85	12.03	41.69	0.092
1976	0.96	1.08	0.69	64.17	32	28.22	12.19	43.19	0.007
1977	1.06	1.19	0.65	54.56	34	28.55	12.34	43.22	0.063
1978	1.19	1.35	0.84	62.38	34	28.79	12.23	42.49	0.022
1979	1.32	1.50	0.86	57.27	34	28.69	12.34	43.02	0.090
1980	1.59	1.81	1.04	57.30	33	28.54	12.18	42.66	0.075
1981	1.80	2.13	1.77	82.95	32	28.28	12.19	43.10	0.045
1982	1.93	2.27	1.27	55.74	32	27.69	12.34	44.57	0.119
1983	2.08	2.50	1.49	59.83	32	27.92	12.59	45.10	0.075
1984	2.20	2.62	1.56	59.55	33	28.14	12.51	44.46	0.089
1985	2.37	2.86	2.10	73.49	33	28.19	12.69	45.00	0.050

Table 2b Female Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	SKEWNESS GRWAGE	SKEWNESS HOURS
1970	2053	8.91	9.33	5.31	56.91	0.667	1.39	4.16	-0.55
1971	3198	9.86	10.27	5.93	57.74	0.702	1.77	4.07	-0.62
1972	3165	11.12	11.58	6.50	56.15	0.704	1.02	2.47	-0.45
1973	3298	12.33	13.07	7.37	56.36	0.613	1.24	3.45	-0.42
1974	3140	15.04	15.61	8.61	55.06	0.541	1.46	25.52	-0.50
1975	3428	18.52	19.49	10.98	56.33	0.708	1.65	3.02	-0.37
1976	3356	21.00	22.35	12.97	58.05	0.565	1.95	9.44	-0.41
1977	3447	24.90	25.23	13.72	54.38	0.612	0.95	5.27	-0.38
1978	3303	28.47	29.53	16.22	54.92	0.572	0.89	7.80	-0.34
1979	3259	32.58	33.70	18.94	56.22	0.577	1.01	8.42	-0.35
1980	3421	38.74	40.34	23.35	57.89	0.650	1.52	7.58	-0.37
1981	3538	42.14	45.64	28.16	61.70	0.473	1.91	24.60	-0.35
1982	3361	44.35	47.82	29.21	61.09	0.698	1.15	3.48	-0.37
1983	2996	48.84	52.27	32.59	62.35	0.660	1.53	3.74	-0.22
1984	3228	52.50	55.84	33.48	59.96	0.631	1.28	5.50	-0.31
1985	3100	56.00	60.63	36.75	60.61	0.568	1.33	15.17	-0.25

Table 3a Female Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.29	0.30	0.10	33.26	30	27.28	13.17	47.39	0.053
1971	0.33	0.34	0.13	37.88	30	27.78	13.23	47.62	0.051
1972	0.37	0.38	0.13	34.59	29	26.87	13.09	48.72	0.135
1973	0.44	0.45	0.17	36.92	26	26.43	12.92	48.88	0.065
1974	0.54	0.56	0.24	42.43	25	26.03	12.71	48.82	0.087
1975	0.74	0.73	0.26	35.02	25	26.16	12.75	48.75	0.162
1976	0.90	0.90	0.37	41.06	25	25.73	13.01	50.55	0.093
1977	0.97	0.98	0.30	30.45	25	25.85	13.16	50.90	0.199
1978	1.07	1.10	0.49	44.87	25	26.20	13.20	50.41	0.125
1979	1.17	1.22	0.39	31.76	25	26.22	13.47	51.37	0.223
1980	1.39	1.44	0.47	32.35	25	25.81	12.92	50.08	0.247
1981	1.54	1.62	0.55	38.67	25	25.70	12.92	50.26	0.243
1982	1.66	1.73	0.59	33.40	23	24.59	13.10	53.29	0.234
1983	1.75	1.85	0.66	35.58	22	24.66	13.25	53.73	0.232
1984	1.88	1.97	0.71	36.03	24	25.01	13.44	53.74	0.144
1985	2.00	2.08	0.77	37.05	24	25.30	13.89	54.88	0.168

Table 3b Female Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	GRWAGE	HOURS
1970	975	7.19	7.32	3.87	52.96	0.551	0.45	1.40	-0.19
1971	1468	8.15	8.15	4.34	53.31	0.494	0.42	3.80	-0.29
1972	1366	9.00	8.98	5.00	55.71	0.603	0.49	1.39	-0.18
1973	1521	10.08	10.43	5.69	54.58	0.501	0.57	2.25	-0.04
1974	1346	12.50	12.42	6.62	53.29	0.462	0.36	6.90	-0.14
1975	1437	15.04	15.55	8.33	53.57	0.610	0.46	1.08	-0.07
1976	1510	17.38	18.11	10.67	58.90	0.596	3.20	5.48	-0.10
1977	1496	19.46	20.04	10.70	53.40	0.591	0.41	0.80	-0.03
1978	1384	22.39	23.32	12.85	55.10	0.519	0.49	13.31	0.01
1979	1364	25.15	26.46	14.81	55.99	0.663	0.57	0.76	0.19
1980	1350	29.46	31.09	17.26	55.52	0.654	0.61	1.09	-0.03
1981	1392	31.50	34.00	19.31	56.77	0.700	0.84	1.39	0.03
1982	1336	32.00	34.67	20.15	58.11	0.615	0.76	1.60	0.09
1983	1062	34.15	37.49	22.56	60.17	0.674	0.86	1.58	0.10
1984	1211	36.75	40.22	23.11	57.45	0.514	0.74	3.60	0.11
1985	1139	41.56	43.22	26.16	60.52	0.623	1.07	2.53	0.21

Table 4a Female Non-Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.39	0.46	0.28	62.42	36	31.97	10.73	33.56	-0.132
1971	0.43	0.49	0.30	61.03	36	31.52	10.46	33.18	-0.073
1972	0.49	0.55	0.30	55.20	35	31.47	10.97	34.87	-0.109
1973	0.54	0.64	0.38	59.75	35	30.99	11.07	35.72	-0.149
1974	0.67	0.77	0.65	84.94	35	30.97	10.74	34.68	-0.054
1975	0.87	0.99	0.54	54.11	35	30.79	11.08	35.98	-0.001
1976	1.04	1.22	0.84	69.00	35	30.26	11.07	36.59	-0.098
1977	1.19	1.35	0.78	58.04	35	30.62	11.25	36.27	-0.054
1978	1.31	1.53	0.98	64.25	35	30.65	11.11	36.25	-0.092
1979	1.51	1.70	1.03	60.56	35	30.46	11.13	36.53	-0.004
1980	1.77	2.06	1.22	59.44	35	30.32	11.32	37.32	-0.036
1981	2.08	2.46	2.16	87.93	35	29.95	11.38	38.00	-0.037
1982	2.24	2.63	1.46	55.43	35	29.73	11.36	38.21	0.004
1983	2.43	2.85	1.69	59.33	35	29.70	11.84	39.87	-0.035
1984	2.59	3.01	1.78	59.33	35	30.02	11.52	38.37	0.001
1985	2.84	3.32	2.47	74.36	35	29.88	11.62	38.88	-0.035

Table 4b Female Non-Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	SKEWNESS GRWAGE	SKEWNESS HOURS
1970	1078	10.58	11.15	5.76	51.62	0.661	1.45	3.49	-0.90
1971	1730	11.36	12.08	6.48	53.64	0.725	1.95	3.48	-0.92
1972	1799	12.96	13.55	6.81	50.25	0.699	1.03	2.02	-0.60
1973	1777	14.36	15.34	7.87	51.31	0.608	1.30	2.94	-0.77
1974	1794	17.45	18.04	9.13	50.58	0.548	1.68	22.97	-0.78
1975	1991	21.48	22.34	11.76	52.63	0.709	1.84	2.73	-0.57
1976	1846	24.36	25.81	13.65	52.86	0.537	1.51	8.68	-0.78
1977	1951	28.50	29.21	14.43	49.41	0.591	0.95	4.72	-0.64
1978	1919	32.57	34.01	16.91	49.72	0.549	0.88	6.81	-0.59
1979	1895	37.60	38.90	19.86	51.04	0.532	1.00	7.85	-0.83
1980	2071	44.42	46.37	24.79	53.46	0.625	1.61	7.16	-0.59
1981	2146	50.06	53.18	30.35	57.07	0.422	1.95	21.69	-0.62
1982	2025	52.64	56.50	30.98	54.83	0.673	1.00	3.13	-0.68
1983	1934	56.67	60.39	34.36	56.90	0.622	1.52	3.42	-0.36
1984	2017	61.48	65.22	35.21	53.98	0.604	1.21	5.21	-0.55
1985	1961	67.44	70.74	38.20	54.00	0.527	1.29	14.34	-0.53

Table 5a Male Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (1,w)
1970	0.57	0.64	0.52	81.12	42	43.43	9.68	22.30	-0.114
1971	0.63	0.71	0.40	55.83	40	42.52	9.20	21.63	-0.054
1972	0.71	0.79	0.42	52.35	40	42.74	9.18	21.48	-0.068
1973	0.82	0.92	0.58	62.30	40	42.75	9.18	21.47	-0.107
1974	0.97	1.07	0.56	52.28	40	42.87	9.85	22.98	-0.042
1975	1.24	1.35	0.65	47.75	40	42.68	9.52	22.30	-0.068
1976	1.41	1.54	0.72	46.84	40	42.39	9.29	21.91	-0.077
1977	1.57	1.71	0.78	45.29	40	42.55	9.38	22.03	-0.110
1978	1.78	1.96	0.95	48.45	40	42.34	9.28	21.91	-0.132
1979	2.03	2.22	1.04	46.86	40	42.85	9.85	22.98	-0.097
1980	2.45	2.72	1.38	50.78	40	42.25	10.00	23.66	-0.067
1981	2.81	3.17	1.74	54.76	40	41.67	9.50	22.80	-0.103
1982	3.07	3.40	1.77	52.13	40	41.52	9.55	22.99	-0.094
1983	3.27	3.72	2.14	57.50	40	41.55	9.96	23.98	-0.104
1984	3.45	3.95	2.39	60.56	40	41.55	10.26	24.69	-0.106
1985	3.72	4.32	2.68	62.14	40	41.89	10.58	25.25	-0.104

Table 5b Male Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	GRWAGE	HOURS
1970	3317	20.00	21.01	9.88	47.03	0.586	1.66	24.83	-0.27
1971	5109	21.42	22.83	11.67	51.08	0.896	3.73	3.92	-0.37
1972	5003	24.45	25.84	12.25	47.40	0.704	2.43	2.92	0.03
1973	4946	27.80	29.67	17.19	57.94	0.824	12.22	6.72	-0.26
1974	4481	31.92	33.90	16.54	48.77	0.872	3.52	3.81	0.09
1975	4906	38.84	40.69	17.57	43.18	0.843	1.79	2.68	0.19
1976	4781	43.00	45.19	19.03	42.11	0.837	1.41	2.51	0.08
1977	4737	48.23	50.74	20.97	41.33	0.838	1.54	2.16	0.27
1978	4502	55.50	58.85	24.55	41.71	0.830	1.68	3.11	0.49
1979	4379	64.76	67.92	29.15	42.92	0.849	1.65	2.23	0.23
1980	4425	77.50	82.61	38.17	46.20	0.874	1.81	2.81	0.38
1981	4565	85.00	92.09	43.86	47.63	0.823	2.07	3.89	0.29
1982	4413	92.00	97.66	45.45	46.54	0.847	2.30	3.28	-0.02
1983	3837	96.46	105.03	52.80	50.27	0.807	2.16	4.06	0.20
1984	3861	102.43	110.36	52.27	47.36	0.741	1.70	7.52	0.12
1985	3842	111.92	122.00	65.36	53.57	0.762	3.20	4.95	0.28

Table 6a Male Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (1,w)
1970	0.53	0.54	0.20	36.22	44	44.68	9.95	22.26	0.065
1971	0.58	0.60	0.22	36.24	42	43.50	9.41	21.63	0.109
1972	0.65	0.67	0.24	36.39	42	43.75	9.35	21.36	0.068
1973	0.75	0.78	0.27	34.67	42	43.84	9.42	21.49	0.047
1974	0.89	0.91	0.30	33.17	42	44.01	10.09	22.93	0.106
1975	1.15	1.18	0.41	34.57	42	43.88	9.45	21.54	0.060
1976	1.32	1.35	0.44	32.35	40	43.43	9.07	20.89	0.065
1977	1.45	1.48	0.49	32.78	41	43.66	9.12	20.89	0.043
1978	1.63	1.69	0.57	33.97	40	43.48	9.20	21.16	0.034
1979	1.88	1.92	0.67	34.91	42	44.06	9.98	22.64	0.076
1980	2.23	2.31	0.87	37.42	41	43.36	9.34	22.92	0.056
1981	2.51	2.63	1.00	38.12	40	42.63	9.50	22.28	0.066
1982	2.71	2.84	1.19	42.05	40	42.55	9.64	22.66	0.012
1983	2.89	2.98	1.15	38.53	41	42.36	9.56	22.56	0.114
1984	3.07	3.18	1.21	38.14	40	42.52	10.32	24.28	0.096
1985	3.25	3.47	1.87	53.91	41	42.80	10.76	25.14	0.001

Table 6b Male Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	GRWAGE	HOURS
1970	2274	19.00	19.01	6.98	36.74	0.815	0.06	0.98	-0.43
1971	3441	20.00	20.33	7.69	37.84	0.837	0.38	0.80	-0.75
1972	3288	23.00	23.05	8.56	37.15	0.834	0.31	0.53	-0.24
1973	3244	26.26	26.49	9.40	35.49	0.811	0.20	0.80	-0.40
1974	2952	30.00	30.39	11.05	36.37	0.799	0.36	0.56	-0.19
1975	3161	37.00	37.40	13.27	35.48	0.784	0.51	1.15	-0.10
1976	3118	41.00	41.55	14.28	34.37	0.795	0.62	0.78	0.11
1977	3071	46.00	46.19	15.79	34.18	0.787	0.66	0.99	-0.20
1978	2906	52.37	53.42	18.49	34.62	0.772	0.51	1.11	-0.11
1979	2747	60.50	61.41	22.60	36.80	0.817	0.74	0.94	-0.03
1980	2688	72.00	73.60	28.53	38.77	0.832	0.87	1.35	-0.08
1981	2744	77.96	80.27	32.47	40.44	0.816	1.83	1.42	-0.05
1982	2686	83.50	85.19	32.84	38.55	0.734	0.66	5.78	-0.35
1983	2139	87.58	89.17	35.25	39.53	0.849	0.88	1.48	-0.54
1984	2260	93.55	95.52	38.63	40.44	0.802	1.18	1.17	-0.27
1985	2232	100.64	104.43	44.71	42.81	0.657	1.51	10.61	-0.04

Table 7a Male Non-Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.75	0.86	0.85	97.82	40	40.71	8.48	20.83	-0.172
1971	0.83	0.94	0.55	58.85	40	40.48	8.38	20.71	-0.080
1972	0.93	1.03	0.55	53.60	40	40.80	8.53	20.91	-0.076
1973	1.02	1.20	0.84	69.92	40	40.67	8.31	20.43	-0.141
1974	1.22	1.38	0.78	56.24	40	40.66	8.98	22.08	-0.042
1975	1.49	1.66	0.85	51.22	39	40.51	9.25	22.84	-0.071
1976	1.73	1.90	0.97	51.18	40	40.45	9.37	23.18	-0.100
1977	1.93	2.13	1.00	46.91	39	40.51	9.50	23.44	-0.137
1978	2.23	2.45	1.25	51.09	39	40.29	9.07	22.51	-0.179
1979	2.53	2.73	1.32	48.19	40	40.80	9.27	22.73	-0.151
1980	3.06	3.36	1.75	52.00	39	40.54	9.85	24.29	-0.076
1981	3.59	3.99	2.22	55.66	38	40.22	9.32	23.17	-0.156
1982	3.92	4.29	2.13	49.74	38	39.92	9.17	22.98	-0.102
1983	4.17	4.65	2.67	57.48	38	40.53	10.37	25.58	-0.177
1984	4.53	5.04	3.12	61.85	38	40.19	10.01	24.91	-0.175
1985	4.87	5.49	3.17	57.48	38	40.63	10.19	25.07	-0.141

Table 7b Male Non-Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	SKEWNESS GRWAGE	SKEWNESS HOURS
1970	1043	22.85	25.36	13.29	52.39	0.515	1.42	17.93	-0.09
1971	1668	25.00	27.98	15.97	57.08	0.910	3.59	3.21	0.51
1972	1715	28.63	31.19	15.92	51.05	0.870	2.39	2.46	0.59
1973	1702	31.86	35.75	25.19	70.47	0.813	10.65	5.31	-0.11
1974	1529	36.92	40.68	22.27	54.74	0.891	3.44	3.11	0.69
1975	1745	43.00	46.65	22.23	47.65	0.865	1.77	2.21	0.76
1976	1663	48.55	52.01	24.24	46.60	0.851	1.17	1.91	0.08
1977	1666	55.06	59.11	26.13	44.21	0.848	1.40	1.64	1.12
1978	1596	64.94	68.73	30.43	44.27	0.843	1.65	2.66	1.68
1979	1632	74.56	78.89	35.07	44.45	0.853	1.64	1.84	0.70
1980	1737	89.64	96.55	46.17	47.82	0.883	1.71	2.45	1.03
1981	1821	102.23	109.88	52.04	47.36	0.815	1.79	3.56	0.84
1982	1727	110.77	117.07	54.63	46.67	0.873	2.46	2.36	0.54
1983	1698	114.26	125.01	63.40	50.72	0.764	1.96	3.70	0.98
1984	1601	123.32	131.30	61.12	46.55	0.691	1.51	7.20	0.69
1985	1615	130.05	146.29	80.00	54.69	0.766	3.16	3.45	0.72

Table 8a Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (l,w)
1970	0.45	0.47	0.21	43.57	40	39.61	13.46	33.98	0.347
1971	0.50	0.52	0.23	43.50	40	38.80	12.89	33.22	0.349
1972	0.56	0.58	0.25	43.48	40	38.79	13.08	33.71	0.360
1973	0.64	0.67	0.28	42.24	40	38.28	13.40	35.01	0.354
1974	0.77	0.80	0.33	40.71	40	38.38	13.79	35.92	0.367
1975	1.00	1.04	0.42	40.55	40	38.34	13.41	34.96	0.359
1976	1.16	1.20	0.47	38.75	40	37.66	13.40	33.58	0.331
1977	1.26	1.32	0.50	37.58	40	37.83	13.51	35.71	0.356
1978	1.43	1.50	0.61	41.04	40	37.90	13.37	35.28	0.319
1979	1.60	1.68	0.68	40.16	40	38.14	14.04	36.82	0.368
1980	1.90	2.02	0.86	42.57	40	37.49	13.79	36.78	0.336
1981	2.14	2.29	1.00	43.51	40	36.93	13.42	36.33	0.359
1982	2.30	2.47	1.15	46.71	40	36.58	13.80	37.74	0.319
1983	2.46	2.60	1.15	43.95	40	36.49	13.74	37.64	0.379
1984	2.56	2.76	1.21	43.90	39	36.41	14.21	39.04	0.353
1985	2.73	3.00	1.72	57.29	39	36.89	14.50	39.31	0.244

Table 8b Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	GRWAGE	HOURS
1970	3249	15.79	15.50	8.21	52.96	0.852	0.25	1.01	-0.67
1971	4909	16.88	16.69	8.84	53.00	0.852	0.41	0.99	-0.85
1972	4654	19.27	18.92	10.01	52.90	0.863	0.34	0.72	-0.69
1973	4765	21.84	21.36	11.25	52.66	0.836	0.29	0.90	-0.65
1974	4298	25.00	24.76	12.93	52.20	0.807	0.38	1.27	-0.53
1975	4598	31.00	30.57	15.66	51.24	0.821	0.41	1.08	-0.58
1976	4628	34.69	33.90	17.18	50.69	0.801	0.57	1.41	-0.58
1977	4567	38.40	37.63	18.86	50.13	0.815	0.43	1.01	-0.65
1978	4290	44.47	43.71	21.98	50.28	0.777	0.37	2.89	-0.61
1979	4111	50.00	49.81	26.17	52.54	0.844	0.55	1.02	-0.45
1980	4038	59.50	59.38	32.31	54.40	0.852	0.70	1.38	-0.47
1981	4136	64.90	64.70	36.09	55.78	0.848	1.22	1.44	-0.51
1982	4022	69.00	68.41	37.70	55.11	0.780	0.57	4.79	-0.58
1983	3201	73.00	72.03	39.89	55.38	0.861	0.69	1.47	-0.63
1984	3471	75.73	76.23	43.04	56.46	0.817	0.90	1.40	-0.47
1985	3371	83.00	83.75	48.92	58.41	0.710	1.14	9.64	-0.36

Table 9a Non-Manual Workers

Normal hourly gross wage (£), normal weekly hours

YEAR	MEDIAN GRWAGE	MEAN GRWAGE	STD GRWAGE	CV in% GRWAGE	MEDIAN HOURS	MEAN HOURS	STD HOURS	CV in% HOURS	COR (1,w)
1970	0.52	0.66	0.66	100.43	38	36.27	10.63	29.30	0.012
1971	0.59	0.71	0.50	69.56	38	35.92	10.50	29.23	0.134
1972	0.66	0.78	0.50	64.07	38	36.02	10.90	30.27	0.137
1973	0.75	0.91	0.71	77.28	38	35.73	10.94	30.63	0.072
1974	0.88	1.05	0.77	73.84	38	35.43	11.07	31.26	0.133
1975	1.11	1.31	0.78	59.48	38	35.33	11.35	32.14	0.153
1976	1.30	1.54	0.97	62.69	38	35.09	11.49	32.74	0.072
1977	1.47	1.71	0.97	56.83	38	35.18	11.58	32.91	0.095
1978	1.67	1.95	1.20	61.84	37	35.03	11.30	32.27	0.054
1979	1.88	2.18	1.28	58.76	37	35.25	11.53	32.70	0.121
1980	2.28	2.65	1.62	61.13	37	34.98	11.82	33.79	0.127
1981	2.67	3.16	2.32	73.32	37	34.66	11.67	33.66	0.072
1982	2.94	3.39	1.98	58.38	37	34.42	11.58	33.65	0.144
1983	3.12	3.69	2.38	64.48	37	34.77	12.41	35.71	0.074
1984	3.31	3.91	2.66	68.14	37	34.52	11.99	34.74	0.084
1985	3.59	4.30	3.00	69.81	37	34.73	12.23	35.20	0.087

Table 9b Non-Manual Workers

Sample sizes, normal weekly net earnings (£), normal hourly gross wage (£), normal weekly hours

YEAR	SAMPLE SIZE	MEDIAN NEARN	MEAN NEARN	STD NEARN	CV in% NEARN	COR (b,w)	SKEWNESS NEARN	SKEWNESS GRWAGE	SKEWNESS HOURS
1970	2121	14.75	18.14	12.41	68.44	0.593	1.73	19.42	-0.73
1971	3398	16.51	19.88	14.48	72.85	0.896	3.29	3.18	-0.56
1972	3514	18.52	22.16	15.01	67.72	0.871	2.26	2.41	-0.43
1973	3479	21.00	25.32	21.12	83.41	0.819	9.80	5.28	-0.73
1974	3323	24.06	28.46	20.01	70.31	0.794	3.16	8.98	-0.38
1975	3736	29.20	33.70	21.25	63.07	0.853	1.81	2.31	-0.26
1976	3509	33.05	38.23	23.40	61.20	0.758	1.40	4.08	-0.51
1977	3617	38.09	42.98	25.47	59.26	0.793	1.45	2.41	-0.19
1978	3515	44.30	49.78	29.59	59.44	0.773	1.58	3.62	-0.10
1979	3527	51.00	57.41	34.33	59.79	0.782	1.56	3.38	-0.41
1980	3808	60.62	69.26	43.94	63.45	0.829	1.76	3.43	-0.14
1981	3967	69.97	79.21	50.39	63.62	0.683	1.75	10.98	-0.31
1982	3752	75.10	84.38	52.94	62.74	0.841	2.00	2.42	-0.43
1983	3632	79.17	90.60	59.56	65.73	0.767	1.93	3.53	-0.03
1984	3618	84.23	94.46	58.49	61.92	0.722	1.56	6.51	-0.22
1985	3576	92.87	104.86	71.43	68.12	0.724	2.90	6.42	-0.18

References

Abbott, M. and O. Ashenfelter (1976), "Labour Supply, Commodity Demand and the Allocation of Time", *Review of Economic Studies*, **43**: 389-411.

---------- (1979), "Labour Supply, Commodity Demand and the Allocation of Time: Correction", *Review of Economic Studies*, **46**: 567-569.

Aitchison, J. and J.A.C. Brown (1957), *The Lognormal Distribution*, Cambridge: Cambridge University Press.

Akerlof, G.A. (1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, **84**: 488-500.

---------- (1982), "Labor Contracts as Partial Gift Exchange", *Quarterly Journal of Economics*, **97**: 543-569.

Akerlof, G.A. and J.L. Yellen (eds.) (1986), *Efficiency Wage Models of the Labor Market*, Cambridge: Cambridge University Press.

Apostol, T.M. (1974), *Mathematical Analysis*, Massachusetts: Addison-Wesley Publishing Company.

Arrow, K.J. and G. Debreu (1954), "Existence of an Equilibrium for a Competitive Economy", *Econometrica*, **22**: 265-290.

Arrow, K.J. and F.H. Hahn (1971), *General Competitive Analysis*, San Francisco: Holden-Day.

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

---------- (ed.) (1976), *The Personal Distribution of Incomes*, London: George Allen & Unwin Ltd.

---------- (1979), "Personelle Einkommens- und Vermögensverteilung", in *Handwörterbuch der Mathematischen Wirtschaftswissenschaften*, Band 1, ed. R. Selten, pp. 315-322, Wiesbaden: Gabler.

---------- (1984), *The Economics of Inequality*, 2nd edition (reprint), Oxford: Oxford University Press.

---------- (1987), "On the Measurement of Poverty", *Econometrica*, **55**: 749-764.

Atkinson, A.B. and J. Micklewright (1983), "On the Reliability of Income Data in the Family Expenditure Survey 1970-1977", *Journal of the Royal Statistical Society*, **146**: 33-61.

---------- (1992), *Economic Transformation in Eastern Europe and the Distribution of Income*, Cambridge: Cambridge University Press.

Atkinson, A.B., J. Micklewright and N.H. Stern (1988), "A Comparison of the Family Expenditure Survey and the New Earnings Survey 1971-1977", Chapter 7 in *Tax-Benefit Models*, eds. A.B. Atkinson and H. Sutherland, pp. 154-222, London: Suntory-Toyota International Centre for Economics and Related Disciplines (London School of Economics).

Atkinson, A.B. and N.H. Stern (1981) (in conjunction with J. Gomulka), "On Labour Supply and Commodity Demands", Chapter 10 in *Essays in the Theory and Measurement of Consumer Behaviour*, ed. A. Deaton, pp. 265-296, Cambridge: Cambridge University Press.

Atkinson, A.B. and J.E. Stiglitz (1980), *Lectures on Public Economics*, London: McGraw-Hill.

Aumann, R.J. (1964), "Markets with a Continuum of Traders", *Econometrica*, **32**: 39-50.

Barten, A.P. and V. Böhm (1981), "Consumer Theory", Chapter 9 in *Handbook of Mathematical Economics*, Vol. II, eds. K.J. Arrow and M.D. Intriligator, pp. 381-429, New York: North Holland.

Becker, G.S. (1985), "Changes in Labor Market Discrimination over Time", *Journal of Labor Economics*, **3** (supplement): S33-S59.

Beckmann, M.J. (1977), "Management Production Functions and the Theory of the Firm", *Journal of Economic Theory*, **14**: 1-18.

Blanchard, O.J. and S. Fischer (1989), *Lectures on Macroeconomics*, Cambridge: MIT Press.

Blinder, A.S. (1973), "Wage Discrimination: Reduced Form and Structural Variation", *Journal of Human Resources*, **8**: 436-455.

---------- (1980), "The Level and Distribution of Economic Well-Being", Chapter 6 in *The American Economy in Transition*, ed. M. Feldstein, Chicago: University of Chicago Press.

Blundell, R. and I. Walker (1982), "Modelling the Joint Determination of Household Labour Supplies and Commodity Demands", *Economic Journal*, **92**: 351-364.

---------- (1986), "A Life-Cycle Consistent Empirical Model of Family Labour Supply using Cross-Section Data", *Review of Economic Studies*, **53**: 539-558.

Browning, M. and C. Meghir (1989), "Testing for the Separability of Commodity Demands from Male and Female Labour Supply", IFS-Working Paper Series, No. 89/4, Institute of Fiscal Studies, London.

Burden, R.L., J.D. Faires and A.C. Reynolds (1981), *Numerical Analysis*, Boston, Mass.: Prindle, Weber and Schmidt.

Calvo, G.A. and S. Wellisz (1978), "Supervision, Loss of Control, and the Optimum Size of the Firm", *Journal of Political Economy*, **86**: 943-952.

---------- (1979), "Hierarchy, Ability, and Income Distribution", *Journal of Political Economy*, **87**: 991-1010.

Chiappori, P.A. (1985), "Distribution of Income and the 'Law of Demand'", *Econometrica*, **53**: 109-127.

Chipman, J., L. Hurwicz, M. Richter and H. Sonnenschein (eds.) (1971), *Preferences, Utility and Demand*, New York: Harcourt-Brace-Javonovich.

Coase, R.H. (1937), "The Nature of the Firm", *Economica*, **4**: 386-405.

Cochran, W.G. (1952), "The x^2 Test of Goodness of Fit", *Annals of Mathematical Statistics*, **23**: 315-345.

Corcoran, M. and G.J. Duncan (1979), "Work History: Labour Force Attachment and Earnings Differences between the Races and Sexes", *Journal of Human Resources*, **14**: 1-20.

Courant, R. and F. John (1974), *Introduction to Calculus and Analysis*, Vol. II, New York: John Wiley & Sons.

Cournot, A. (1838), *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Hachette: Paris. Translated as *Researches into the Mathematical Principles of the Theory of Wealth*, 1929, New York: McMillan.

Darling, D.A. (1957), "The Kolmogorov-Smirnov, Cramér-von Mises Tests", *Annals of Mathematical Statistics*, **28**: 823-838.

Deaton, A. (1988), "Agricultural Pricing Policies and Demand Patterns in Thailand", Discussion Paper No. 136, Woodrow Wilson School, Princeton University.

Deaton, A. and J. Muellbauer (1980), "An Almost Ideal Demand System", *American Economic Review*, **70**: 312-326.

Debreu, G. (1959), *Theory of Value*, New Haven and London: Yale University Press. (Originally published by John Wiley & Sohns.)

---------- (1970), "Economies with a Finite Set of Equilibria", *Econometrica*, **38**: 387-392.

---------- (1974), "Excess Demand Functions", *Journal of Mathematical Economics*, **1**: 15-21.

---------- (1981), "Existence of Competitive Equilibrium", Chapter 15 in *Handbook of Mathematical Economics*, Vol. II, eds. K.J. Arrow and M.D. Intriligator, pp. 697-743, New York: North Holland.

Department of Health and Social Security (1983), *Tables on Families with Low Incomes*, London.

Dierker, E. (1974), *Topological Methods in Walrasian Economics*, Berlin: Springer-Verlag.

Duncan, G.J. and S. Hoffman (1978), "On-the-Job Training and Earnings Differences by Race and Sex", *Review of Economics and Statistics*, **61**: 594-602.

Eubank, R.L. (1988), *Spline Smoothing and Nonparametric Regression*, New York and Basel: Marcel Dekker.

Freixas, X. and A. Mas-Colell (1987), "Engel Curves Leading to the Weak Axiom in the Aggregate", *Econometrica*, **55**: 515-531.

Friedman, M. (1953), "Choice, Chance, and the Personal Distribution of Income", *Journal of Political Economy*, **61**: 277-290.

Gale, D. (1963), "A Note on Global Instability of Competitive Equilibrium", *Naval Research Logistics Quarterly*, **10**: 81-87.

Gasser, T. and H.G. Müller (1979), "Kernel Estimation of Regression Functions", in *Smoothing Techniques for Curve Estimation*, eds. T. Gasser and H. Rosenblatt, pp. 23-68, Berlin: Springer-Verlag.

Gibrat, R. (1931), *Les Inégalités Economiques*, Paris: Recveil Sirey.

Gösecke, G. and K.D. Bedau (1974), "Verteilung und Schichtung der Einkommen der privaten Haushalte in der Bundesrepublik Deutschland 1950-1975", in *Deutsches Institut für Wirtschaftsforschung*, Heft 31, Berlin: Duncker und Humbolt.

Gomulka, J. and N. Stern (1990), "The Employment of Married Women in the United Kingdom 1970-83", *Economica*, **57**: 171-199.

Grandmont, J.-M. (1987), "Distributions of Preferences and the 'Law of Demand'", *Econometrica*, **55**: 155-161.

---------- (1992), "Transformations of the Commodity Space, Behavioral Heterogeneity, and the Aggregation Problem", *Journal of Economic Theory*, **57**: 1-35.

Gunderson, M. (1989), "Male-Female Wage Differentials and Policy Responses", *Journal of Economic Literature*, **27**: 46-72.

Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge: Cambridge University Press.

Härdle, W. and M. Jerison (1988), "Cross Section Engel Curves over Time", Discussion Paper No. A-160, Sonderforschungsbereich 303, Projektbereich A, University of Bonn.

Harrison, A. (1981), "Earnings by Size: A Tale of Two Distributions", *Review of Economic Studies*, **48**: 621-631.

Hildenbrand, K. and W. Hildenbrand (1986), "On the Mean Income Effect: A Data Analysis of the U.K. Family Expenditure Survey", Chapter 14 in *Contributions to Mathematical Economics*, eds. W. Hildenbrand and A. Mas-Colell, pp. 247-268, Amsterdam: North Holland.

Hildenbrand, K. and S. Islam (1985), "Models of Income Distribution in Britain: Lognormal, Gamma and Beta Densities", Discussion Paper No. A-16, Sonderforschungsbereich 303, Projektbereich A, University of Bonn.

Hildenbrand, W. (1974), *Core and Equilibria of a Large Economy*, Princeton: Princeton University Press.

---------- (1983), "On the 'Law of Demand'", *Econometrica*, **51**: 997-1020.

---------- (1985a), "A Problem in Demand Aggregation: Per Capita Demand as a Function of Per Capita Expenditure", Discussion Paper No. A-12, Sonderforschungsbereich 303, Projektbereich A, University of Bonn.

---------- (1985b), "Über den empirischen Gehalt der neoklassischen ökonomischen Theorie", in *Rheinisch-Westfälische Akademie der Wissenschaften, Vorträge N 336*, pp. 36-58: Westdeutscher Verlag.

---------- (1989a), "Facts and Ideas in Microeconomic Theory", *European Economic Review*, **33**: 251-276.

---------- (1989b), "The Weak Axiom of Revealed Preference for Market Demand is Strong", *Econometrica*, **57**: 979-985.

Hildenbrand, W. and M. Jerison (1988), "The Demand Theory of the Weak Axioms of Revealed Preference", Discussion Paper No. A-163, Sonderforschungsbereich 303, Projektbereich A, University of Bonn.

Hirsch, M.W. and S. Smale (1974), *Differential Equations, Dynamical Systems, and Linear Algebra*, New York: Academic Press.

Hogarth, R.M. and M.W. Reder (eds.) (1987), *Rational Choice: The Contrast between Economics and Psychology*, Chicago and London: Chicago University Press.

Holmström, B.R. and J. Tirole (1989), "The Theory of the Firm", Chapter 2 in *Handbook of Industrial Organization*, Vol. I, eds. R. Schmalensee and R.D. Willig, pp. 61-133, Amsterdam: North Holland.

Houthakker, H.S. (1950), "Revealed Preference and the Utility Function", *Economica*, **17**: 159-174.

Hurwicz, L. and H. Uzawa (1971), "On the Integrability of Demand Functions", in *Preferences, Utility and Demand*, eds. J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein, New York: Harcourt-Brace-Javonovich.

Johnson, N.L. and S. Kotz (1970a), *Continuous Univariate Distributions-1*, New York: John Wiley & Sons.

---------- (1970b), *Continuous Univariate Distributions-2*, Boston: Houghton Mifflin Company.

Kaiser, H. (1990), "On the Estimation of a Commodity Demand and Labour Supply System for West Germany", Discussion Paper No. TIDI/144 (Taxation, Incentives and the Distribution of Income Programme), Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics.

Kannai, Y. (1987), "A Characterization of Monotone Individual Demand Functions", Discussion Paper No. A-101, Sonderforschungsbereich 303, Projektbereich A, University of Bonn.

Kehoe, T.J. (1985), "Multiplicity of Equilibria and Comparative Statics", *Quarterly Journal of Economics*, **100**: 119-148.

Kemsley, W.F.F. (1975), "Family Expenditure Survey: A Study of Differential Response based on a Comparison of the 1971 Sample with the Census", *Statistical News*.

Kemsley, W.F.F., R.U. Redpath and M. Holmes (1980), *Family Expenditure Survey Handbook*, London: Her Majesty's Stationary Office.

Kendall, M.G. and A. Stuart (1973), *The Advanced Theory of Statistics*, Vol. II, 3rd edition, London: Griffin.

Kihlstrom, R., A. Mas-Colell and H. Sonnenschein (1976), "The Demand Theory of the Weak Axiom of Revealed Preference", *Econometrica*, **44**: 971-978.

Killingsworth, M.R. (1983), *Labor Supply*, Cambridge: Cambridge University Press.

Killingsworth, M.R. and J. Heckman (1986), "Female Labor Supply", Chapter 2 in *Handbook of Labor Economics*, Vol. I, eds. O. Ashenfelter and R. Layard, pp. 103-204, Amsterdam: North Holland.

Kirman, A.P. and K.-J. Koch (1986), "Market Excess Demand in Exchange Economies with Identical Preferences and Collinear Endowments", *Review of Economic Studies*, **53**: 457-463.

Koopmans, T.J. (1957), *Three Essays on the State of Economic Science*, New York: McGraw-Hill.

Krishnaiah, P.R. and P.K. Sen (eds.) (1984), *Handbook of Statistics*, Vol 4, Nonparametric Methods, Amsterdam: North Holland.

Laha, R.G. and V.K. Rohatgi (1979), *Probability Theory*, New York: John Wiley & Sons.

Loève, M. (1977), *Probability Theory*, Vol. I, New York-Heidelberg-Berlin: Springer-Verlag. (Republication of the 1963 editition, D. Van Nostrand)

Lucas, R.E. and L. Rapping (1970), "Real Wages, Employment and Inflation", in *Microeconomic Foundations of Employment and Inflation Theory*, eds. E.S. Phelps et al., pp. 257-305, New York: Norton.

Lydall, H.F. (1959), "The Distribution of Employment Incomes", *Econometrica*, **27**: 110-115.

---------- (1968), *The Structure of Earnings*, Oxford: Oxford University Press.

Malcomson, J.M. (1984), "Work Incentives, Hierarchy, and Internal Labor Markets", *Journal of Political Economy*, **92**: 486-507.

Malinvaud, E. (1972), *Lectures on Microeconomic Theory*, Amsterdam: North Holland.

Mantel, R.R. (1976), "Homothetic Preferences and Community Excess Demand Functions", *Journal of Economic Theory*, **12**: 197-201.

Mas-Colell, A. (1985), *The Theory of General Economic Equilibrium: A Differentiable Approach*, Cambridge: Cambridge University Press.

Massey, F.J. (1950), "A Note on the Power of a Non-Parametric Test", *Annals of Mathematical Statistics*, **20**: 440-442.

---------- (1952), "Correction to 'A Note on the Power of a Non-Parametric Test'", *Annals of Mathematical Statistics*, **23**: 637-638.

Mayer, T. (1960), "The Distribution of Ability and Earnings", *Review of Economics and Statistics*, **42**: 189-195.

Medoff, J.L. and K.G. Abraham (1980), "Experience, Performance, and Earnings", *Quarterly Journal of Economics*, **95**: 703-736.

Mirrlees, J.A. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**: 175-208.

---------- (1976), "The Optimal Structure of Incentives and Authority within an Organization", *Bell Journal of Economics*, **7**: 105-131.

Mitjuschin, L.G. and W.M. Polterovich (1978), "Criteria for Monotonicity of Demand Functions" (in Russian), *Ekonomika i Matématicheski Metody*, **14**: 122-128.

Ortega, J.M. and W.C. Rheinbold (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, New York: Academic Press.

Parzen, E. (1962), "On Estimation of a Probability Density Function and Mode", *Annals of Mathematical Statistics*, **33**: 1065-1076.

Pencavel, J. (1986), "Labor Supply of Men", Chapter 1 in *Handbook of Labor Economics*, Vol. I, eds. O. Ashenfelter and R. Layard, pp. 3-102, Amsterdam: North Holland.

Prakasa Rao, B.L.S. (1983), *Nonparametric Functional Estimation*, New York: Academic Press.

Reinsch, C.H. (1967), "Smoothing by Spline Functions", *Numerische Mathematik*, **10**: 177-183.

Rosen, S. (1982), "Authority, Control, and the Distribution of Earnings", *Bell Journal of Economics*, **13**: 311-323.

Routh, G. (1980), *Occupation and Pay in Great Britain 1906-1979*, London: MacMillan Press.

Sahota, G.S. (1978), "Theories of Personal Income Distribution: A Survey", *Journal of Economic Literature*, **16**: 1-55.

Samuelson, P.A. (1938), "A Note on the Pure Theory of Consumer's Behavior", *Economica*, **5**: 61-71.

---------- (1947), *Foundations of Economic Analysis*, Cambridge, Massachusetts: Harvard University Press.

Scarf, H. (1960), "Some Examples of Global Instability of the Competitive Equilibrium", *International Economic Review*, **1**: 157-172.

Schmalensee, R. and R.D. Willig (eds.) (1989), *Handbook of Industrial Organization*, Vol. I, New York: North Holland.

Schultz, T.P. (1980), "Estimating Labor Supply Functions for Married Women", Chapter 1 in *Female Labor Supply: Theory and Estimation*, ed. J.P. Smith, pp. 25-89, Princeton: Princeton University Press.

Scott, D.W., R.A. Tapia and J.R. Thompson (1980), "Nonparametric Probability Density Estimation by Discrete Maximum Penalized-Likelihood Criteria", *Annals of Statistics*, **8**: 820-832.

Shafer, W. and H. Sonnenschein (1981), "Market Demand and Excess Demand Functions", Chapter 14 in *Handbook of Mathematical Economics*, Vol. II, eds. K.J. Arrow and M.D. Intriligator, pp. 671-693, New York: North Holland.

Shorrocks, A.F. (1983), "Ranking Income Distributions", *Economica*, **50**: 3-17.

Silverman, B.W. (1985), "Some Aspects of the Spline Smoothing Approach to Nonparametric Regression Curve Fitting (with Discussion)", *Journal of the Royal Statistical Society Series B*, **46**: 1-52.

---------- (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Simon, H. (1955), "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, **69**: 99-118.

---------- (1956), "Rational Choice and the Structure of the Environment", *Psychological Review*, **63**: 129-138.

---------- (1957), "The Compensation of Executives", *Sociometry*, **20**: 32-35.

---------- (1972), "Theories of Bounded Rationality", in *Decision and Organization*, eds. C.B. McGuire and R. Radner, pp. 161-176, Amsterdam: North Holland.

---------- (1976), "From Substantive to Procedural Rationality", in *Methods and Appraisal in Economics*, ed. S.J. Latis, pp. 129-148, Cambridge: Cambridge University Press.

Smith, J.P. (ed.) (1980), *Female Labor Supply: Theory and Estimation*, Princeton: Princeton University Press.

Sonnenschein, H. (1973), "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?", *Journal of Economic Theory*, **6**: 345-354.

Stark, T. (1978), "Personal Incomes", in *Reviews of United Kingdom Statistical Sources*, Vol. VI, Oxford: Pergamon.

Stephens, M.A. (1974), "EDF Statistics for Goodness of Fit and Some Comparisons", *Journal of the American Statistical Association*, **69**: 730-737.

Stern, N. (1986), "On the Specification of Labour Supply Functions", in *Unemployment, Search and Labour Supply*, eds. R. Blundell and I. Walker, Cambridge: Cambridge University Press.

Stiglitz, J.E. (1975), "Incentives, Risk, and Information: Notes towards a Theory of Hierarchy", *Bell Journal of Economics*, **6**: 552-579.

Stone, C.H. (1977), "Consistent Nonparametric Regression", *Annals of Statistics*, **5**: 595-645.

Szymanski, S. (1987), "Moral Hazard, Internal Labour Markets and Hierarchy", Discussion Paper No. 88/2, Department of Economics, Birkbeck College, University of London.

Tapia, R.A. and J.R. Thompson (1978), *Nonparametric Probability Density Estimation*, Baltimore: Johns Hopkins University Press.

Thaler, R.H. (1989), "Anomalies - Interindustry Wage Differentials", *Journal of Economic Perspectives*, **3**: 181-193.

Thurow, L. (1976), *Generating Inequality*, London: Macmillan.

Tirole, J. (1988), *The Theory of Industrial Organization*, Cambridge: MIT Press.

Trockel, W. (1983), *Market Demand: An Analysis of Large Economies with Non-Convex Preferences*, Berlin: Springer-Verlag.

Tuck, R.H. (1954), *An Essay on the Economic Theory of Rank*, Oxford: Basil Blackwell.

Tukey, J.W., (1961), "Curves as Parameters and Touch Estimation", Proceedings of the 4th Berkely Symposium: 681-694.

Varian, H.R. (1982), "The Nonparametric Approach to Demand Analysis", *Econometrica*, **50**: 945-973.

---------- (1983), "Non-Parametric Tests of Consumer Behaviour", *Review of Economic Studies*, **50**: 99-110.

---------- (1984), *Microeconomic Analysis*, 2nd edition, New York and London: W.W. Norton and Company.

Vives, X. (1987), "Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand", *Review of Economic Studies*, **54**: 87-103.

Wald, A. (1935), "Über die eindeutige positive Lösbarkeit der neuen Produktionsgleichungen", *Ergebnisse eines mathematischen Kolloquiums*, ed. K. Menger, No. 6: 12-20.

---------- (1936a), "Über die Produktionsgleichungen der ökonomischen Wertlehre", *Ergebnisse eines mathematischen Kolloquiums*, ed. K. Menger, No. 7: 1-6.

---------- (1936b), "Über einige Gleichungssysteme der mathematischen Ökonomie", Zeitschrift für Nationalökonomie, **7**: 637-670. Translated as "On some Systems of Equations of Mathematical Economics", 1951, Econometrica, **19**: 368-403.

Waldman, M. (1984), "Worker Allocation, Hierarchies and the Wage Distribution", *Review of Economic Studies*, **51**: 95-109.

Walras, L. (1874), *Eléments d'économie politique pure*, Corbaz: Lausanne. English translation by W. Jaffé, 1954, *Elements of Pure Economics*, London: Allen and Unwin.

von Weizsäcker, R.K. (1987), *Theorie der Verteilung der Arbeitseinkommen*, Tübingen: Mohr.

Williamson, O.E. (1967), "Hierarchical Control and Optimum Firm Size", *Journal of Political Economy*, **75**: 123-138.

Zabalza, A. and Z. Tsannatos (1985), *Women and Equal Pay: the Effects of Legislation on Female Employment in Britain*, Cambridge: Cambridge University Press.