Community Choice and Local Income Taxation

Kurt Schmidheiny

Inauguraldissertation

zur Erlangung der Würde eines

Doctor rerum oeconomicarum

der Wirtschafts- und Sozialwissenschaftlichen

Fakultät der Universität Bern

Bern, August 2003

Preface

Zollikon near Zurich, Riehen near Basel, Cologny near Geneva, Muri near Bern, Saint-Sulpice near Lausanne, Teufen near St. Gallen and Hergiswil near Lucerne are the pleasant-sounding names of the tax h(e)avens that accompany every major Swiss city. The residents of these posh communities have far higher incomes and pay taxes at substantially lower local rates than their neighbors in the city. Why is it that not all households wish to live in such a place?

This thesis studies the phenomenon of spatial income segregation in systems of communities with local income taxation. The various chapters explore the location of households and the corresponding politico-economic equilibrium of tax rates and housing prices from both a theoretical and an empirical perspective. The theoretical reflections are based on microeconomic models with heterogeneous agents. The empirical investigation follows a microeconometric approach using discrete choice models. The research touches on the fields of urban public finance, fiscal federalism and public choice.

A central feature of my thesis is the extensive use of computationally intensive numerical methods. Both the simulations in the theoretical models and the estimations in the empirical investigation require numerical multidimensional integration routines and numerical optimization algorithms.

Chapter 1 (Kapitel 1) is a German summary of the theoretical concepts, models and simulation studies in Chapters 2 and 3. It also presents the descriptive results of the empirical investigation in Chapter 4.

Income segregation in theoretical models of urban public finance is analyzed in Chapter 2. I develop a set of necessary conditions for segregation of the population in models with housing markets and heterogeneous households. The conditions for the sorting of the population according to income classes or other dimensions of heterogeneity are established without explicitly describing the household utility function and budget constraint. They therefore apply to a broad class of models, including models with income taxation and property taxation. The segregation conditions in the existing literature are surveyed using a common framework, and a series of new and less specific models is proposed. The analysis shows that in models with income taxation, segregation can often only be established under very restrictive assumptions on the household's preferences.

Chapter 3 presents a model of an urban area with local income taxation used to finance a local public good. Households are allowed to differ in both incomes and tastes for housing. In a calibrated two-community model, I show the existence of an equilibrium which features segregation of households by both incomes and tastes. The high-tax community shows lower housing prices and lower public good provision than the low-tax community. The model is able to explain the substantial differences in local income tax levels and average income across communities as observed in Swiss metropolitan areas. The numerical investigation suggests that taste heterogeneity reduces the distributional effects of local tax differences. The numerical investigation also suggests that the ability of a rich community to set low taxes is higher when this community is physically small.

Swiss metropolitan areas are comprised of a system of communities with considerable fiscal autonomy and thus provide an excellent setting in which to assess the empirical relevance of multi-community models. Chapter 4 investigates how the income tax differentials across communities in an urban area affect the households' location decisions. I first develop a theoretical model with progressive income taxation that explains income segregation in the institutional setting of Switzerland. The empirical investigation uses data from the urban agglomeration of Basel for the year 1997. This unique data set contains tax information from all households that moved either within the city center of Basel or from the city center to the outskirts. The community choice of the households is investigated within the framework of a random utility maximization model which is derived from the theoretical model. I compare different econometric specifications of the error term structure, such as conditional logit, nested logit and multinomial probit. The empirical results show that rich households are significantly

and substantially more likely to move to low-tax communities than poor households.

Acknowledgement

My thanks go first to my supervisor Klaus Neusser for his support and his mathematical expertise. Christian Keuschnigg from the Universtity of St. Gallen kindly agreed to act as a co-referee. I am deeply indebted to Veronika Aegerter, my most relentless critic. I would like to thank Frank Bodmer, Esther Brügger, Marius Brühlhart, Harris Dellas, Winand Emons, Dennis Epple, Lars Feld, Mike Gerfin, Werner Hildenbrand, Anke Kessler, Robi Leu, Christoph Lülfesmann, Kenneth Train, Manuel Wälti, Gabi Wanzenried for valuable comments and suggestions. I have greatly benefited from the Brown-Bag seminars at the Economics Department in Bern and the Crans-Montana Workshop on Globalisation and Economic Geography. I am grateful to the Tax Administration and the Statistical Office of the Canton of Basel-Stadt for the provision of the unique household data set, my thanks to Martin Sandtner in particular. Thanks are also due to Wüest und Partner, and especially to Jürg Mosimann for the provision of the housing price data. The hospitality of the HWWA Hamburg Institute of International Economics is gratefully acknowledged. Adrian Herzog's free MAPresso software was used to produce the informative maps. Heather Murray helped me find the right words.

Dank

Meinen Eltern Helen und Hans danke für Ihre Unterstützung auf die ich immer zählen konnte. Die volkswirtschaftlichen Diskussionen mit meinem Papa sind mir eine stete Herausforderung.

Contents

List of Figures

List of Tables

Kapitel 1

Wohnsitzwahl und lokale Einkommensbesteuerung – Eine Einführung

1.1 Einleitung

Fiskalischer Föderalismus wird in der Europäischen Union heftig debattiert. Einerseits möchten die EU Finanzminister 'schädlichen' Steuerwettbewerb unterbinden, andererseits legt die gewünschte Dezentralisierung von Entscheidungsprozessen nach dem Subsidiaritätsprinzip regionale Steuerautonomie nahe.

Die Vorteile der dezentralen Entscheidungsfindung *und* Finanzierung von lokalen Angelegenheiten sind offensichtlich: Die betroffene Bevölkerung kennt ihre eigenen Bedürfnisse am Besten und wird den Nutzen öffentlicher Projekte optimal gegen die Kosten abwägen (Oates, 1972). Ausserdem trägt die Konkurrenzsituation zwischen den Gemeinden dazu bei, staatliche Aktivität möglichst effizient zu organisieren. Als stossend empfinden jedoch viele die daraus resultierende Ungerechtigkeit durch z.T häufig grosse Unterschiede in der Steuerbelastung. Oft wird auch argumentiert, dass die staatliche Fähigkeit, Steuern zu erheben, erodiere, wenn mobile Haushalte in steuergünstige Gebiete abwandern können.

Das Studium der lokalen Versorgung mit lokalen öffentlichen Gütern geht auf Tiebout (1956) zurück. Tiebout zeigte, dass fiskalischer Föderalismus zu einer effizienten Produktion von lokalen öffentlichen Gütern führt, weil sich Haushalte mit gleichen Präferenzen in einer Gemeinde gruppieren und das ihnen am besten passende Angebot an öffentlichen Gütern bereit stellen. Zentral bei Tiebouts Modell ist, dass alle Haushalte über die gleichen Ressourcen verfügen. Wenn sich die Haushalte in ihrem Einkommen unterscheiden, ist weder die Existenz eines Gleichgewichts noch die Effizienz desselben garantiert. In einer solchen Situation ist es für reiche Haushalte vorteilhaft, sich mit anderen reichen Haushalten in der gleichen Gemeinde niederzulassen und dadurch die Steuerbelastung gering zu halten, ohne auf ein hohes Niveau an lokalen öffentlichen Leistungen verzichten zu müssen. Eine so geformte Gemeinde wird dabei grundsätzlich auch für ärmere Haushalte attraktiv, die von den Steuereinnahmen der reichen Haushalte profitieren möchten. Während sich also reiche Haushalte in reichen Gemeinden zu gruppieren versuchen, sind die armen Haushalte bestrebt, ihnen nachzufolgen. Sie hinterlassen so immer kleinere und immer ärmere Gemeinden. Wie die Übersicht über die Literatur im nächsten Abschnitt aufzeigt, muss dieser negative Kreislauf jedoch nicht zwangsläufig einsetzen.

Tiebouts Analyse löste eine lange und fortwährende Diskussion zu fiskalischem Föderalismus auf Gemeindeebene aus. Ein Meilenstein war die Berücksichtigung von heterogenen Haushaltseinkommen durch Ellickson (1971) und Westfhoff (1977). Mit der Modellierung von heterogenen Haushalten verschob sich das Forschungsinteresse von der Suche nach optimalen Gemeindegrössen und -grenzen hin zur Untersuchung von Gemeindesystemen mit festen politischen Grenzen. Diese Modelltradition konzentriert sich auf die Analyse von Modellen mit einer lokalen *'property tax'*. Die lokale Steuer auf Bodenbesitz und Wohneigentum ist in den Vereinigten Staaten die hauptsächliche lokale Steuerart. Die umfangreiche theoretische und empirische Literatur wird von Ross und Yinger (1999) ausführlich dargestellt. Einen sehr umfassenden Überblick bietet auch Feld (1999). Besonders hervorzuheben sind die Modelle von Epple et al. (1984, 1993), Epple und Romer (1991) sowie Epple und Platt (1998), denen die Modelle in diesem Beitrag formal am nächsten sind.

Dieser Beitrag richtet den Blick auf lokale *Einkommenssteuern*, wie sie die Gemeinden in Metropolregionen der Schweiz erheben. Modelle mit lokalen Einkommenssteuern und Einkommensheterogenität wurden in der bisherigen Literatur nur vereinzelt diskutiert. Goodspeed (1989), Calabrese (1990) sowie Hansen und Kessler (2001a) untersuchen ähnliche Modelle wie das hier in Abschnitt 1.4 vorgestellte, jedoch unter restriktiveren und teilweise sehr unrealistischen Annahmen.

Eine Gemeinsamkeit der erwähnten Modelle liegt in der Analyse von Gleichgewichten, in denen sich die Haushalte aus verschiedenen Einkommensschichten in verschiedenen Gemeinden gruppieren. Die Entstehung des Phänomens der *Einkommenssegregation* wird im Folgenden kurz skizziert. Damit die verschiedenen Gemeinden für ihre Bewohner attraktiv sind, müssen sie negative Standortfaktoren wie hohe Steuern durch positive Standortfaktoren wie tiefe Wohnpreise oder eine gutes Angebot an öffentlichen Gütern kompensieren. Diese Spiegelung von Unterschieden der lokalen Steuerbelastung und des lokalen Angebots an öffentlichen Gütern in den Boden- und damit Wohnpreisen wird als *Kapitalisierung* der Standortvorteile in den Bodenpreisen bezeichnet. Die Kompensation im Gleichgewicht wird nicht von allen Haushalten gleich bewertet. Reichere Haushalte werden etwa durch niedrige Wohnpreise weniger für hohe Steuern kompensiert als ärmere Haushalte, was die unterschiedliche Wohnsitzwahl von Haushalten aus unterschiedlichen Einkommensklassen erklärt. Diese unterschiedliche Abwägung der Gemeindecharakteristika kann vielfältige Gründe haben, wie die beiden Modelle in den Abschnitten 1.3 und 1.4 zeigen. Die Abhängigkeit zwischen Einkommensverteilung und Steuerbelastung ist in diesen Modellen eine grundsätzlich *wechselseitige*: Eine reiche Wohnbevölkerung generiert grosse Einnahmen und ermöglicht tiefe Steuern, während tiefe Steuern besonders reiche Haushalte anziehen.

Abschnitt 1.2 dieses Kapitels stellt das Grundmodell einer Metropolregion mit lokalen Einkommenssteuern und heterogenen Haushaltseinkommen vor und beleuchtet die Bedingungen, unter denen Gleichgewichte mit Einkommenssegregation entstehen. Abschnitt 1.3 zeigt, wie *progressive* Einkommenssteuern Einkommenssegregation auslösen. In Abschnitt 1.4 wird ein politökonomisches Modell analysiert, in dem Einkommenssegregation durch die sinkende Bedeutung von Wohnausgaben im Budget von reichen Haushalten entsteht. Die simulierten Gleichgewichte in den kalibrierten Modellen vermögen grosse und stabile lokale Steuer- und Einkommensunterschiede zu erklären. Beide Modelle erreichen durch die zusätzliche Berücksichtigung von heterogenen Präferenzen eine realistischere Beschreibung städtischer Metropolregionen. Die Modellergebnisse werden in Abschnitt 1.5 für zwei schweizerische Metropolregionen empirisch überprüft.

1.2 Das Grundmodell

Das theoretische Modell beschreibt eine städtische Metropolregion mit einer festen Anzahl Gemeinden, deren politische Grenzen als gegeben angenommen werden. Die Metropolregion wird von heterogenen Haushalten bevölkert, die innerhalb der Metropolregion ihre Wohnsitzgemeinde frei wählen.[1]

In dieser vereinfachten Ökonomie gibt es drei Güter: privaten Konsum, das Gut Wohnen und ein lokales öffentliches Gut. Wohnen meint den Konsum von Boden in Form einer Wohnung oder eines Hauses. Zur Vereinfachung wird angenommen, dass alle Haushalte Mieter ihres Wohnraums sind und die Grundstückseigentümer 'ausserhalb des Modells' leben. Das private Gut fasst alle weiteren von einem Haushalt konsumierten Güter zusammen.

Jede Gemeinde produziert und finanziert ein lokales öffentliches Gut. Das öffentliche Gut steht dabei für ein von den einzelnen Gemeinden produziertes und finanziertes Gut und weist nicht zwingend alle Charakteristika eines reinen öffentlichen Gutes auf.[2] Das Gut ist öffentlich in dem Sinne, dass es von allen Einwohnern einer Gemeinde im gleichen Ausmass konsumiert wird. Bewohner anderer Gemeinden sind vom Konsum ausgeschlossen. Es umfasst beispielsweise Schulen, Abfallentsorgung, öffentliche Sicherheit, Sozialfürsorge, Strassenbau und -unterhalt oder Planungsaufgaben. Finanziert wird das öffentliche Gut aus den Einnahmen einer Einkommenssteuer auf die Wohnbevölkerung. Das Niveau des Angebots an öffentlichen Gütern wird entweder von einer höheren politischen Ebene festgelegt (Abschnitt 1.3) oder in einer Abstimmung durch die Wohnbevölkerung bestimmt (Abschnitt 1.4).

[1] Als technische Vereinfachung wird die Bevölkerung als ein durch eine Dichtefunktion beschriebenes Kontinuum von Haushalten beschrieben.

[2] Reine öffentliche Güter zeichnen sich durch Nichtrivalität und Nichtausschliessbarkeit im Konsum aus. Dies bedeutet, dass niemand vom Konsum des Gutes ausgeschlossen werden kann und der Konsum durch weitere Individuen den Nutzen für andere Konsumenten nicht reduziert. Das klassische Beispiel eines reinen öffentlichen Gutes ist der Leuchtturm. Hingegen erfüllen z.B. ein öffentliches Schwimmbad oder öffentliche Schulen beide Bedingungen nicht.

Zu jeder Gemeinde gehört eine für den Wohnungsbau verfügbare Fläche, die durch politische Grenzen definiert ist. Das Angebot an Wohnraum wächst mit der Zunahme der bebaubaren Landfläche und mit steigendem Wohnpreis. Der positive Zusammenhang erklärt sich einerseits durch zunehmende Bautätigkeit und andererseits durch eine Verdichung der bestehenden Bauten bei höheren lokalen Haus- und Wohnungspreisen. Der Wohnpreis in jeder Gemeinde ergibt sich aus lokalem Angebot an und lokaler Nachfrage nach Wohnraum.

Jede Gemeinde lässt sich beschreiben durch das lokale Angebot an öffentlichen Gütern, durch ihren Einkommenssteuersatz und ihr Wohnpreisniveau. Dieses Bündel von Gemeindecharakteristika bewerten die Haushalte bei der Wahl ihres Wohnsitzes.

Haushalte werden durch ihr Einkommen und ihre Präferenzen beschrieben. Sie unterscheiden sich in ihrem Einkommen und, in einer Erweiterung des Modells, in ihren Präferenzen. Ein Haushalt wählt diejenige Gemeinde, in der er den grössten Nutzen aus privatem und öffentlichem Konsum erwartet. Hat ein Haushalt eine Wohnsitzgemeinde gewählt, so bestimmt er seinen Konsum an Wohnraum und anderen privaten Gütern, indem er unter Berücksichtigung seines Haushaltsbudgets seinen Nutzen maximiert.[3]

1.2.1 Asymmetrische Gleichgewichte und Einkommenssegregation

Die oben beschriebene Metropolregion befindet sich im Gleichgewicht, wenn der Wohnungsmarkt in allen Gemeinden im Gleichgewicht ist, das Angebot an öffentlichen Gütern durch einen dezentralen politischen Prozess oder exogen bestimmt wurde, das Gemeindebudget in allen Gemeinden ausgeglichen ist und kein Haushalt die Gemeinde wechseln möchte. Der Wohnsitzentscheid der Haushalte legt die mögliche räumliche Struktur des Gleichgewichtszustandes weitgehend fest.

Es lassen sich zwei Typen von Gleichgewichten unterscheiden: *Symmetrische Gleichgewichte*, in denen die Haushalte indifferent sind zwischen den Gemeinden und sich zufällig niederlassen und *asymmetrische Gleichgewichte*, in denen Haushalte aus unterschiedlichen Einkommensklassen un-

[3] Der private Konsum dient als Numeraire und sein Preis wird auf 1 gesetzt, ohne dass die Allgemeinheit der Aussage eingeschränkt wird.

terschiedliche Gemeinden vorziehen. In symmetrischen Gleichgewichten ist die Einkommensverteilung in allen Gemeinden gleich. Alle Gemeinden weisen das gleiche Niveau an Steuersätzen und Wohnpreisen auf und verfügen über das gleiche Angebot an öffentlichen Gütern. In asymmetrischen Gleichgewichten sind Haushalte nach Einkommensklassen *segregiert*. Dieser zweite Typ von Gleichgewichten wird im Folgenden den Schwerpunkt bilden, weil wir in städtischen Metropolregionen mit lokalen Einkommenssteuern üblicherweise Einkommenssegregation beobachten und weil symmetrische Gleichgewichte häufig nicht stabil sind.[4]

Die Existenz eines Gleichgewichts und seine qualitativen Eigenschaften hängen auch vom Wohnungsmarkt und von der Ausgestaltung des politischen Prozesses ab. In den folgenden Abschnitten wird untersucht, ob und unter welchen Bedingungen in der beschriebenen Metropolregion ein asymmetrisches Gleichgewicht entstehen kann und welche Eigenschaften dieses aufweist.

1.2.2 Bedingungen für Einkommenssegregation

Dieser Abschnitt zeigt, dass Einkommenssegregation nur unter spezifischen Annahmen über die Präferenzen der Haushalte entsteht. Als Beispiel dient ein Modell mit den in der ökonomischen Modellbildung üblichen homothetischen Haushaltspräferenzen und linearen Einkommenssteuern.

Die Präferenzen des Haushalts werden zur Vereinfachung der Darstellung durch die einfachste Form von homothetischen Präferenzen, durch eine Cobb-Douglas Nutzenfunktion, beschrieben. Das Ergebnis gilt jedoch für alle homothetischen Nutzenfunktionen wie in Kapitel 4 gezeigt wird. Der Haushalt bestimmt seinen Konsum an Wohnraum und anderen privaten Gütern so, dass sein Nutzen maximiert wird. Dabei berücksichtigt er sein Haushaltsbudget. Der Konsumentscheid eines Haushalts kann formal wie folgt dargestellt werden:

$$\max_{h,b} \; U(h,b,g) = h^{\alpha} b^{1-\alpha} g^{\gamma} \quad N.B. \quad ph + b \leq y(1-t)\,,$$

[4] Instabilität bedeutet hier, dass z.B. der Ortswechsel eines einzelnen reichen Haushalts einen Migrationsprozess auslöst, der in ein asymmetrisches Gleichgewicht führt. In den in der Literatur und hier behandelten statischen Modellen kann die Frage der Stabilität aber nur unter ad-hoc Annahmen untersucht werden.

wobei h der Konsum an Wohnraum, b an privatem Gut und g an öffentlichem Gut ist und y für das Haushaltseinkommen, p für den (Miet-)Preis von Wohnraum und t für den proportionalen Einkommenssteuersatz steht. Homothetische Präferenzen zeichnen sich dadurch aus, dass die Grenzrate der Substitution zwischen den Gütern vom Nutzenniveau unabhängig ist. Die Grenzrate der Substitution meint das Austauschverhältnis zwischen beispielsweise dem privaten Gut b und dem öffentlichen Gut g, bei dem der Nutzen des Haushalts unverändert bleibt. Bei homothetischen Präferenzen bleibt dieses Austauschverhältnis bei einer Vervielfachung des Konsums aller Güter gleich.

Wie bewertet nun ein Haushalt Gemeinden mit verschiedenen Steuersätzen, Wohnpreisen und unterschiedlichem Angebot an öffentlichen Gütern? Jeder Haushalt bevorzugt, ceteris paribus, Gemeinden mit tiefen Steuern, tiefen Preisen und gutem Angebot an öffentlichen Gütern. In einem asymmetrischen Gleichgewicht, in dem alle Gemeinden besiedelt sind, werden deshalb einzelne Gemeinden in der einen Dimension, z.B. durch tiefe Steuern, attraktiv sein, während andere Gemeinden in einer anderen Dimension, z.B. mit tiefen Wohnpreisen, vorteilhaft erscheinen. Die Haushalte müssen zwischen den vorteilhaften und den nachteiligen Gemeindecharakteristika abwägen. Für jeden Haushalt gibt es ein Austauschverhältnis (eine Grenzrate der Substitution von Gemeindecharakteristika) von z.B. Steuersatz und Wohnpreisniveau, bei dem er gerade indifferent zwischen zwei Gemeinden ist. Bei anderen Austauschverhältnissen wählt der Haushalt die für ihn vorteilhafte Gemeinde.

Die Annahme homothetischer Präferenzen in Verbindung mit proportionalen Einkommenssteuern hat zur Folge, dass dieses Austauschverhältnis unabhängig vom Einkommen eines Haushalts ist (vgl. die Beweisführung im Anhang). Das Haushaltseinkommen bestimmt zwar selbstverständlich das Nutzenniveau eines Haushalts; ein reicher und ein armer Haushalt wägen jedoch tiefe Steuern genau gleich gegen hohe Wohnpreise oder gegen ein schlechteres Angebot an öffentlichen Gütern ab. Dies hat zur Folge, dass entweder alle Haushalte dieselbe Gemeinde bevorzugen oder alle Haushalte zwischen den Gemeinden indifferent sind. Unter diesen (Standard-)Bedingungen kommt es also zu keiner räumlichen Gruppierung von bestimmten Einkommensklassen:

Satz 1 (Nichtexistenz Einkommenssegregation)
Wenn die Gemeinden proportionale Einkommenssteuern erheben und die Präferenzen der Haushalte homothetisch sind, so existieren keine Gleichgewichte, in denen reiche Haushalte andere Gemeinden bevorzugen als arme Haushalte.

In Kapitel 2 werden die notwendigen und hinreichenden Bedingungen gezeigt, die im dargestellten Grundmodell Einkommenssegregation auslösen. In den folgenden Abschnitten werden zwei realitätsnahe Annahmen vorgestellt, die zu Einkommenssegregation führen.

1.3 Einkommenssegregation durch Steuerprogression

Dieser Abschnitt stellt eine Variante des Grundmodells vor, in dem die Progressivität der lokalen Einkommenssteuern zu einem Gleichgewicht mit Einkommenssegregation führt. Der *Steuertarif*, der die Form der Steuerprogression festlegt, ist für alle Gemeinden der Metropolregion gleich. Die einzelnen Gemeinden legen jedoch individuell ihren *Steuerfuss* (Hebesatz) fest, der den Steuertarif proportional verschiebt. So gestaltet sich der lokale Steuerwettbewerb innerhalb des gleichen Kantons in den grossen Schweizer Metropolregionen.[5] Im Anhang wird dieses Modell formal dargestellt und werden die Ergebnisse hergeleitet. In Kapitel 4 wird das gleiche Modell mit weniger strikten Annahmen an die Nutzenfunktion diskutiert.

Das Angebot an öffentlichen Gütern wird hier von einer höheren politischen Ebene festgelegt und ist in allen Gemeinden gleich. In der Schweiz können die Gemeinden ihre Ausgaben prinzipiell autonom bestimmen. Die Gemeindeautonomie wird aber durch gesetzliche Regelungen auf höherer Ebene zum Teil stark eingeschränkt. So sind z.B. Gemeinden nicht frei, die Grösse von Schulklassen oder die Löhne von Lehrerinnen und Lehrern festzulegen. Viele von der höheren Ebene festgelegte Ausgaben werden jedoch weitgehend durch lokal erhobene Steuern finanziert. Die Annahme exogen

[5]Eine Ausnahme bildet der Kanton Basel-Stadt, dessen Gemeinden Riehen und Bettingen bis 2001 eigene Steuertarife festlegen konnten.

festgelegter, aber lokal finanzierter öffentlicher Ausgaben trägt diesen Einschränkungen Rechnung.

Wie im Beispiel in Abschnitt 1.2.2 werden die homothetischen Präferenzen der Haushalte durch eine Cobb-Douglas Nutzenfunktion beschrieben.[6] Die Haushalte unterscheiden sich vorerst nur in ihrem Einkommen y.[7] Der Konsumentscheid eines Haushalts kann formal wie folgt dargestellt werden:

$$\max_{h,b} U(h,b,g) = h^{\alpha} b^{1-\alpha} g^{\gamma} \quad N.B. \quad ph + b \leq y[1 - t\, r(y)]\,.$$

Der lokale Steuerfuss t und der Steuertarif $r(y)$ bilden dabei den lokalen einkommensabhängigen Steuersatz $t \cdot r(y)$. Der Steuertarif ist progressiv und garantiert, dass der Durchschnittssteuersatz kleiner ist als der Grenzsteuersatz und beide nicht über 100 % liegen.

Da das Angebot an öffentlichen Gütern in allen Gemeinden identisch ist, vergleicht ein Haushalt bei der Wahl seines Wohnsitzes nur das Bündel aus lokalem Steuerfuss und lokalem Wohnpreis. In einem Gleichgewicht, in dem alle Gemeinden besiedelt sind, darf keine Gemeinde bezüglich beider Charakteristika unattraktiver sein als die anderen Gemeinden:

Satz 2 (Kapitalisierung im Modell mit Steuerprogression)
Wenn alle Gemeinden besiedelt sind und unterschiedliche Gemeindecharakteristika aufweisen, so erheben Gemeinden mit tieferen Wohnpreisen höhere Steuern.

Satz 2 bedeutet, dass das tiefe Steuerniveau in steuergünstigen Gemeinden in den Wohnpreisen ‘kapitalisiert’ wird.

Die Progressivität der Einkommenssteuern beeinflusst die Wohnsitzwahl der verschiedenen Haushalte entscheidend. Anders als im Beispiel mit proportionalen Steuern bewerten Haushalte mit unterschiedlichem Einkommen die unterschiedlichen Bündel von Gemeindecharakteristika nun unterschiedlich. Für reiche Haushalte ist die absolute Differenz der *Steuersätze*

[6]Homothetische Präferenzen fokussieren auf den Effekt der Steuerprogression. Die Ergebnisse gelten jedoch allgemein, solange der Budgetanteil für Wohnen mit dem Einkommen sinkt. Vgl. Kapitel 4.

[7]Die Bevölkerung setzt sich aus einem Kontinuum von Haushalten mit einem Einkommen zwischen einem Minimum y_{min} und einem Maximum y_{max} zusammen.

grösser als für ärmere Haushalte. Dies hat zur Folge, dass reiche Haushalte dem Steuerfuss mehr Gewicht beimessen und deshalb Gemeinden mit tiefen Steuerfüssen vorziehen (vgl. auch den Beweis im Anhang). Aus diesen Überlegungen folgt:

Satz 3 (Einkommenssegregation durch Steuerprogression)
Wenn alle Gemeinden besiedelt sind und unterschiedliche Gemeindecharakteristika aufweisen, sind alle Einwohner einer Gemeinde mit tieferem Steuerfuss reicher als alle Einwohner einer Gemeinde mit höherem Steuerfuss.

Die ersten beiden Sätze legen die mögliche räumliche Einkommensstruktur des Gemeindesystems fest. Es blieb bisher jedoch offen, ob diese Struktur durch ein Gleichgewicht gestützt wird und welche Eigenschaften dieses Gleichgewicht allenfalls aufweist. Im Folgenden soll deshalb ein vollständig spezifiziertes und kalibriertes Modell untersucht werden.

Zur Vervollständigung des Modells müssen das Angebot an Wohnraum und die Produktion des öffentlichen Gutes festgelegt werden (vgl. die formale Darstellung im Anhang). Wie im Grundmodell beschrieben, ist das Wohnraumangebot eine zunehmende Funktion der Landfläche der Gemeinde und des lokalen Wohnpreises. Die Kosten für die Produktion des öffentlichen Gutes nehmen linear zu mit steigender Einwohnerzahl der Gemeinde. Im kalibrierten Modell werden konstante Skalenerträge angenommen.

Das Modell wird für die Metropolregion Zürich kalibriert. Diese wird durch die Stadt Zürich und alle Gemeinden definiert, aus denen mehr als 1/3 der arbeitstätigen Bevölkerung in die Zentrumsgemeinde pendeln. Die 41 Gemeinden umfassen eine bebaubare Fläche von $140km^2$, wovon $53km^2$ auf die Stadt Zürich entfallen. 1998 betrug die Einwohnerzahl 628'00, 334'000 in der Zentrumsgemeinde Zürich. Die so definierte Metropolregion wird im Modell durch zwei autonome Gemeinden mit einer relativen Landfläche von 0.4 (Zentrum) resp. 0.6 (Peripherie) dargestellt. Der Steuerfuss der Zentrumsgemeinde war 1997 19% höher als der durchschnittliche Steuerfuss der Peripheriegemeinden. Gemäss den zur Kalibrierung verwendeten Daten beträgt das Medianeinkommen im gesamten Gebiet CHF 67'000 und ist in der Stadt Zürich mit CHF 59'000 rund 23% tiefer als in den umliegenden Gemeinden (CHF 75'000). Die Details zur Kalibrierung sind im Anhang beschrieben.

Tabelle 1.1: Gleichgewichtswerte im Modell mit Steuerprogression.

		homogene Präferenzen		heterogene Präferenzen	
	harmonisiert	Zentrum	Peripherie	Zentrum	Peripherie
L: Land	1	0.4	0.6	0.4	0.6
p: Mietpreis	11.7	6.6	13.1	9.9	12.5
t: Steuerfuss	1	5.17	0.91	2.30	0.87
n: Einwohner	1	0.12	0.88	0.23	0.77
Ey: Mittl. Einkommen	78'547	30'771	85'010	47'755	87'703
$Etr(y)$: Mittl. Steuersatz	0.055	0.161	0.053	0.095	0.051

Die kalibrierten Modellparameter: $g = 5000$, $E(ln\,y) = 11.1$, $SD(ln\,y) = 0.55$, $y_{min} = 23'000$, $y_{max} = 500'000$, $E(\alpha) = 0.25$, $S.A.(\alpha) = 0.11$ (heterogene Präferenzen), $S.A.(\alpha) = 0$ (homogene Präferenzen), $\theta = 3$, $c_0 = 0$, $c_1 = 1$, $r_0 = 0.132$, $r_1 = 1$ und $r_2 = 0.00001$.

Im Gleichgewicht ist der Wohnungsmarkt in beiden Gemeinden geräumt, das Budget beider Gemeinden ausgeglichen und kein Haushalt will den Wohnort wechseln. Die Gleichgewichtswerte können durch numerische Optimierungsalgorithmen unter Verwendung von numerischen Integrationsverfahren gefunden werden. Die Spalten 2 und 3 in Tabelle 1.1 zeigen die Gleichgewichtswerte für die zwei Gemeinden im kalibrierten Modell.[8] Zum Vergleich sind in der ersten Spalte von Tabelle 1.1 die Gleichgewichtswerte angegeben, die sich ergeben, wenn die Steuern in den verschiedenen Gemeinden harmonisiert würden. Dies entspricht der Situation mit einer einzigen politischen Gemeinde.

Die beiden Gemeinden weisen im Gleichgewicht sehr grosse Unterschiede auf: Der Steuerfuss der Zentrumsgemeinde ist fast sechs mal höher als derjenige der Peripheriegemeinden, während die Wohnpreise in der Peripherie rund doppelt so hoch sind wie im Zentrum. Da Haushalte mit einem Einkommen von über CHF 41'000 in die steuergünstige Peripherie ziehen und alle ärmeren Haushalte im Zentrum leben, beträgt das Durchschnittseinkommen im Zentrum nur ein Drittel desjenigen in den umliegenden Gemeinden. Reiche Haushalte können die hohen Wohnpreise in der Peripherie durch die tiefen Steuersätze mehr als kompensieren. Ärmere Haushalte wer-

[8] Die Bezeichnungen 'Zentrum' und 'Peripherie' sind im Modell austauschbar.

den dagegen durch die hohen Wohnpreise von einer Übersiedlung in die Peripherie abgehalten. Trotz tieferer Steuersätze kann die reiche Peripheriegemeinde dank des guten Steuersubstrats das öffentliche Gut vollständig finanzieren. Die Einkommenssegregation führt hier zu einer faktischen Umkehr der Steuerprogression, denn die Einwohner der reichen Peripherie sind im Mittel mit tieferen Steuersätzen $Etr(y)$ konfrontiert als die Einwohner des armen Zentrums.

Das vorgestellte Modell vermag die Unterschiede zwischen den Gemeinden qualitativ richtig zu reproduzieren. Es zeigt sich auch, dass eine Situation mit vollständiger Segregation der Einkommen in zwei Gemeinden im Gleichgewicht ist. Die quantitativen Unterschiede der Merkmale zwischen Zentrum und Peripherie sind aber viel grösser als in der Realität beobachtet. Insbesondere entspricht die perfekte Segregation von armen und reichen Haushalten nicht der Realität. Deshalb wird das Modell nun um heterogene Präferenzen erweitert.

1.3.1 Heterogene Präferenzen

Haushalte werden in der Erweiterung des Modells nicht nur bezüglich ihres Einkommens, sondern auch bezüglich ihrer Präferenz für Wohnen als heterogen angenommen.[9] Der Parameter α in der Cobb-Douglas Nutzenfunktion dient der Modellierung von Präferenzheterogenität. α bestimmt das Verhältnis der Wohnausgaben zum Nettoeinkommen. Für einen Haushalt mit einem hohen α nimmt Wohnen einen höheren Stellenwert und damit auch einen grösseren Anteil am Haushaltsbudget ein. Ein Blick auf die zur Kalibrierung verwendeten Daten zeigt eine grosse Heterogenität der relativen Wohnausgaben.

Satz 2 gilt unverändert. Die in Satz 3 propagierte Einkommenssegregation gilt im erweiterten Modell nicht mehr allgemein, sondern nur für Subgruppen der Bevölkerung mit je gleichen Präferenzen:

[9] Die Zusammensetzung der Bevölkerung aus den bezüglich Einkommen y und Wohnpräferenz α unterschiedlichen Haushalten wird durch eine gemeinsame Dichtefunktion $f(y,\alpha)$ beschrieben.

Satz 4 (Bedingte Einkommenssegregation d. Steuerprogression)
Betrachtet wird eine Subgruppe der Bevölkerung mit gleichen Präferenzen. Wenn alle Gemeinden besiedelt sind und unterschiedliche Gemeindecharakteristika aufweisen, so sind alle Einwohner in einer Gemeinde mit tieferem Steuerfuss reicher als alle Einwohner in einer Gemeinde mit höherem Steuerfuss.

Haushalte mit einer grossen Präferenz für Wohnen sind stärker von den Wohnpreisen betroffen als Haushalte, die einen kleineren Einkommensanteil für Wohnen aufwenden. Deshalb wohnen erstere eher in Gemeinden mit tiefen Wohnpreisen (vgl. den formalen Beweis im Anhang). Daraus folgt:

Satz 5 (Bedingte Präferenzsegregation)
Betrachtet wird eine Subgruppe der Bevölkerung mit gleichem Einkommen. Wenn alle Gemeinden unterschiedliche Charakteristika aufweisen und die Haushalte in verschiedenen Gemeinden wohnen, so haben alle Einwohner in einer Gemeinde mit höheren Wohnpreisen eine schwächer Präferenz für Wohnen als alle Einwohner in einer Gemeinde mit tieferen Wohnpreisen.

Die durch die Sätze 4 und 5 bestimmte Aufteilung der Bevölkerung auf die verschiedenen Gemeinden nimmt eine komplexere Form an als im einfachen Modell und lässt sich am besten im kalibrierten Modell erläutern. Vgl. die Details der Kalibrierung im Anhang.

Abbildung 1.1 zeigt die Einkommenssegregation im Gleichgewicht. Die Grafik links zeigt für alle möglichen Kombinationen von Einkommen und Präferenzparameter die bevorzugte Gemeinde. Die Grafik rechts stellt die Einkommensverteilung in den Gemeinden dar. In beiden Gemeinden sind nun reiche wie arme Haushalte anzutreffen. Reiche Haushalte mit einer sehr grossen Präferenz für Wohnen entscheiden sich für das preisgünstige Zentrum, arme Haushalte mit einer geringen Präferenz für Wohnen ziehen in die steuergünstige Peripherie. Der Anteil der reichen Haushalte bleibt in der steuergünstigen Peripherie höher als im Zentrum.

Die Gleichgewichtswerte des kalibrierten Modells mit heterogenen Präferenzen sind in den Spalten 4 und 5 von Tabelle 1.1 dargestellt. Die Werte der Vergleichssituation mit harmonisierten Steuern (Spalte 1) bleiben unverändert. Die Unterschiede zwischen den Gemeinden sind markant kleiner

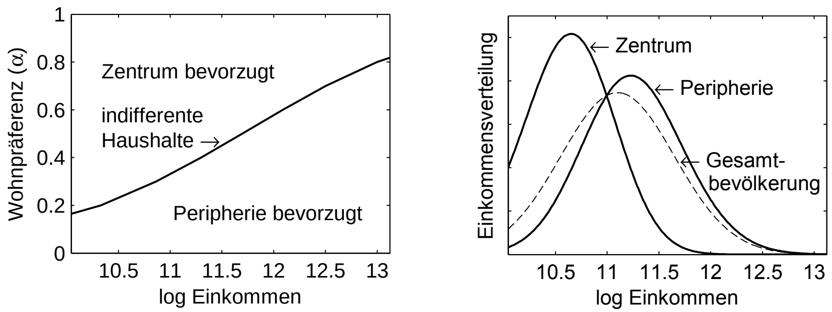

Abbildung 1.1: Einkommens- und Präferenzsegregation im Gleichgewicht des Modells mit progressiven Steuern und heterogenen Präferenzen.

als im Modell ohne Präferenzheterogeniät: Der Steuerfuss im Zentrum ist noch zweieinhalb mal so gross wie derjenige in der Peripherie und auch die Wohnpreis- und Einkommensunterschiede sind geringer. Die Einführung von Präferenzheterogenität liefert deshalb einen wichtigen Beitrag zur quantitativen Erklärung der Situation in Metropolregionen der Schweiz.

1.4 Einkommenssegregation in einem politökonomischen Modell

In diesem Abschnitt wird eine weitere Annahme vorgestellt, die ein asymmetrisches Gleichgewicht unterstützt. Zusätzlich erlaubt dieser Ansatz die Berücksichtigung des lokalen politökonomischen Prozesses zur Bestimmung der lokalen öffentlichen Ausgaben. Das Modell wird detailliert in Kapitel 3 behandelt und soll an dieser Stelle kurz zusammengefasst werden.

Die Präferenzen der Haushalte sind nicht homothetisch und werden durch eine Stone-Geary Nutzenfunktion beschrieben.[10] Die lokalen Einkommenssteuern werden in dieser Variante des Grundmodells wieder als proportional angenommen.[11] Das Optimierungskalkül eines Haushalts in einer

[10]Vgl. Deaton und Muellbauer (1980) für eine ausführliche Besprechung dieser Nutzenfunktion und der daraus resultierenden linearen Nachfragesysteme.

[11]Die Behandlung von nichtlinearen Steuersätzen in Modellen mit endogenem öffentlichem Angebot führt zu grossen technischen Schwierigkeiten. Vgl. dazu Kapitel 2.

gewählten Gemeinde ist:

$$\max_{h,b} U(h,b,g) = (h-\beta_h)^{\alpha}(b-\beta_b)^{1-\alpha}(g-\beta_g)^{\gamma}$$
$$N.B. \qquad ph + b \leq y(1-t)\,,$$

wobei die Parameter β_h, β_b und β_g als existenzsichernder Konsum des Haushalts an Wohnraum, anderen privaten Gütern und an öffentlichem Gut interpretiert werden können. Der Parameter α gilt als Mass für die Wohnpräferenz und liegt in der Variante mit heterogenen Präferenzen zwischen 0 und 1.

Bei seiner Wohnsitzwahl wägt ein Haushalt die Bündel aus Gemeindecharakteristika der verschiedenen Gemeinden gegeneinander ab. Diese Bündel bestehen hier, anders als im Modell mit progressiven Steuern, aus drei Dimensionen: dem Steuersatz, dem Wohnpreis und dem Angebot an öffentlichen Gütern. Alle Haushalte ziehen, ceteris paribus, tiefe Steuern, tiefe Wohnpreise und ein grosses Angebot an öffentlichen Gütern vor. In einem Gleichgewicht, in dem alle Gemeinden besiedelt sind und verschiedene Charakteristika aufweisen, müssen deshalb Gemeinden mit hohen Steuersätzen diese entweder durch tiefe Wohnpreise *oder* durch ein gutes Angebot an öffentlichen Gütern kompensieren. Damit gilt Satz 2 des Modells mit progressiven Steuern nicht mehr allgemein. Es eröffnet sich eine Vielfalt von möglichen Bündeln von Gemeindecharakteristika: Gemeinden mit tiefen Steuern, hohen Wohnpreisen und geringem Angebot an öffentlichen Gütern sind ebenso denkbar wie Gemeinden mit tiefen Steuern, tiefen Wohnpreisen und geringem Angebot an öffentlichen Gütern oder Gemeinden mit tiefen Steuern, teuren Wohnungen und grossem öffentlichem Angebot. Erst das allgemeine Gleichgewicht im kalibrierten Modell zeigt, welche dieser Varianten vorkommt.

In der Stone-Geary Spezifikation hängt das Verhältnis der Wohnausgaben zum Nettoeinkommen vom Einkommen ab. Je nach Präferenzparameter α steigt oder fällt der Budgetanteil des Wohnens mit zunehmendem Einkommen. Anders als im Beispiel mit homothetischen Präferenzen in Abschnitt 1.2.2 sind deshalb die Austauschverhältnisse zwischen den Gemeindecharakteristika auch bei proportionalen Steuern vom Einkommen abhängig. In der Subgruppe der Haushalte mit einer kleinen Wohnpräferenz α wird der Wohnpreis mit steigendem Einkommen immer unbedeutender,

Tabelle 1.2: Gleichgewichtswerte im polit-ökonomischen Modell.

		homogene Präferenzen		heterogene Präferenzen	
	harmonisiert	Zentrum	Peripherie	Zentrum	Peripherie
L: Land	1	0.40	0.60	0.40	0.60
p: Mietpreis	11.4	9.2	12.4	10.5	12.0
t: Steuersatz	0.064	0.110	0.056	0.085	0.059
g: Öffentlich. Gut	5032	4335	5390	4488	5225
n: Einwohner	1	0.31	0.69	0.28	0.72
Ey: Mittleres Einkommen	78'547	39'368	96'460	52,995	88,687

Die kalibrierten Modellparameter: $\beta_h = 700$, $\beta_b = 13000$, $\beta_g = 4000$, $\gamma = 0.02$, $E(\alpha) = 0.17$, $S.A.(\alpha) = 0.11$ (heterogene Präferenzen), $S.A.(\alpha) = 0$ (homogene Präferenzen), $E(ln\, y) = 11.1$, $SD(ln\, y) = 0.55$, $y_{min} = 23'000$, $y_{max} = 500'000$, $\theta = 3$, $c_0 = 0$ und $c_1 = 1$.

während er in der Subgruppe mit hohem α immer wichtiger wird. Die Präferenzsegregation in Satz 5 gilt unverändert. Satz 4 des Modells mit progressiven Steuern gilt in diesem Modell ohne Bindung an die Ordnung der Gemeindecharakteristika:

Satz 6 (Bedingte Einkommenssegregation im politökon. Modell) *Betrachtet wird eine Subgruppe der Bevölkerung mit gleichen Präferenzen. Wenn alle Gemeinden unterschiedliche Charakteristika aufweisen und die Haushalte in verschiedenen Gemeinden wohnen, so gilt für jedes Paar von Gemeinden, dass in der einen Gemeinde entweder alle Einwohner reicher oder alle ärmer sind als in der anderen Gemeinde.*

Nebst der Wohnangebotsfunktion und der Produktionsfunktion muss in dieser Variante der politökonomische Prozess in den Gemeinden modelliert werden. Das Angebot an öffentlichen Gütern wird in einer Abstimmung durch die Wohnbevölkerung bestimmt.[12] Die Stimmenden wissen dabei, dass eine Erhöhung der Ausgaben eine Erhöhung des Steuersatzes mit sich

[12] Abstimmung und Wohnsitzwahl werden in einem zweistufigen Spiel modelliert. Auf der ersten Stufe wählen die Haushalte die Gemeinden. Auf der zweiten Stufe stimmt die Wohnbevölkerung über das öffentliche Angebot ab. Das Ergebnis dieser Abstimmung wird von den Haushalten bei ihrer Wohnsitzwahl auf der ersten Stufe antizipiert.

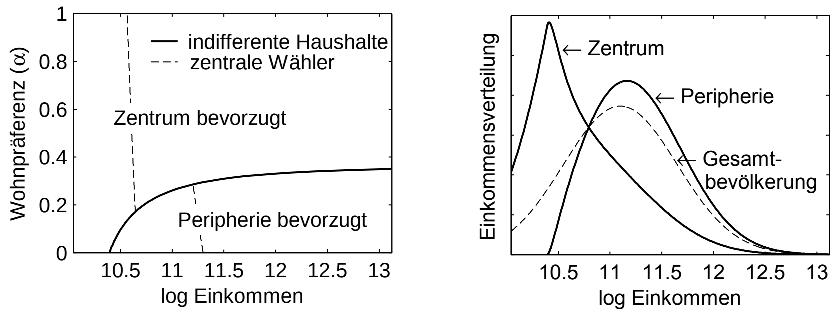

Abbildung 1.2: Einkommens- und Präferenzsegregation im Gleichgewicht des politökonomischen Modells mit heterogenen Präferenzen.

bringt, soll das Gemeindebudget ausgeglichen bleiben. Das öffentliche Angebot und der Steuersatz sind in einem politischen Gleichgewicht, wenn keine andere Kombination existiert, die von mehr als 50% der Bevölkerung bevorzugt wird. In der vereinfachten Modellvariante mit Präferenzheterogenität entspricht dies einem Medianwählermodell.

Damit sind alle Teile des Modells beschrieben. Ein analog zu Abschnitt 1.3 kalibriertes Modell liefert die qualitativen und quantitativen Eigenschaften des Gleichgewichts. Der mit Daten der Schweizerischen Arbeitskräfteerhebung SAKE (1995) geschätzte Parameter α zeigt zum einen eine grosse Heterogenität in der Bevölkerung und zum anderen, dass in der Mehrheit (und im Mittel) der Bevölkerung die relativen Wohnausgaben mit dem Einkommen sinken. Dies ist für die Form der Einkommens- und Präferenzsegregation entscheidend.

Tabelle 1.2 zeigt die Gleichgewichtswerte im kalibrierten Modell. Wie im Modell mit progressiven Steuern werden die Ergebnisse bei homogenen und heterogenen Präferenzen vorgestellt und mit der Situation nach einer Steuerharmonisierung verglichen. Abbildung 1.2 weist die Einkommens- und Präferenzsegregation im Gleichgewicht mit heterogenen Präferenzen aus. Das linke Schaubild zeigt, dass in beiden Gemeinden reiche Haushalte leben. Die ärmsten Haushalte wählen der tiefen Wohnpreise wegen alle die Zentrumsgemeinde. Mit steigendem Einkommen sind Haushalte zunehmend bereit, in die steuergünstige Peripherie zu ziehen. Die sich daraus ergebende

Einkommensverteilung ist im rechten Schaubild dargestellt.

Im Gleichgewicht mit heterogenen Präferenzen (Spalten 4 und 5) erhebt die Gemeinde in der Peripherie 31% tiefere Steuern als die Zentrumsgemeinde und weist einen um 14% höheren Wohnpreis auf. Wie im Modell mit progressiven Steuern ist in der einkommensstärkeren Peripherie das Steuerniveau tiefer als im Zentrum. Während dieser Zusammenhang im Modell mit progressiven Steuern gemäss Satz 2 zwingend entstehen muss, ist er hier ein Ergebnis des allgemeinen Gleichgewichts. Die Differenzen sind im politökonomischen Modell kleiner, weil die reichen Haushalte nicht nur durch tiefe Steuern, sondern auch durch das bessere Angebot an öffentlichen Gütern (16% höher als im Zentrum) von der Peripherie angezogen werden. Andererseits muss die 'arme' Zentrumsgemeinde ihre (proportionalen) Steuern viel weniger anheben, um das Gemeindebudget auszugleichen, als dies im Fall von progressiven Steuern nötig wäre. Wiederum sind die Steuer-, Preis- und Einkommensunterschiede im Modell mit homogenen Präferenzen viel ausgeprägter als wenn heterogene Präferenzen angenommen werden. Die realistische Annahme heterogener Präferenzen verringert also auch in diesem Modell die Gemeindedifferenzen und die damit einhergehenden negativen Konsequenzen der lokalen Steuerhoheit.

1.5 Empirische Evidenz

Die beiden theoretischen Modelle in den Abschnitten 1.3 und 1.4 postulieren zwei empirisch überprüfbare Hypothesen: (1) Die Kapitalisierung von tiefen Steuersätzen in den Wohnpreisen wie in Satz 2 beschrieben und in den numerischen Gleichgewichten der beiden Modelle aufgezeigt. (2) Die Einkommenssegregation gemäss den Sätzen 3, 4 und 6. Die Segregationshypothese hat zwei Aspekte: Einerseits sollte im langfristigen Gleichgewicht eines Gemeindesystems eine Segregation der Einkommen entstehen, die mit den Steuerunterschieden korrespondiert. Andererseits sollte bei der individuellen Wohnsitzwahl eines Haushaltes beobachtet werden können, dass reiche Haushalte systematisch steuergünstige Gemeinden bevorzugen.

Eine Vielzahl von Studien hat das Verhältnis zwischen Steuersätzen, Wohnpreisen und dem Angebot öffentlicher Güter untersucht. Die Kapitalisierung von Steuern und weiteren Standortfaktoren in den Wohnpreisen

wurde von Oates (1969) und in zahlreichen nachfolgenden Untersuchungen für verschiedene Länder deutlich gezeigt. Überzeugend ist diese von Hilber (1998) für die Schweiz nachgewiesen.

Einkommenssegregation im Gleichgewicht wurde von Epple und Sieg (1999) und Epple, Romer und Sieg (2001) für die Metropolregion Boston untersucht. Ihre Schätzungen zeigen, dass die beobachteten Einkommens- und Steuerunterschiede gut durch ein theoretischen Modell in der Tradition von Westhoff und Ellickson erklärt werden können. Rhode and Strumpf (forthcoming) betrachten Segregation in den Vereinigten Staaten aus einer historischen Perspektive und kommen zum Schluss, dass Tiebout-Modelle die beobachtete Entwicklung der Bevölkerungssegregation, insbesondere auf nationaler Ebene, nicht erklären können. Feld und Kirchgässner (2001) zeigen für die Schweizer Kantone und grösseren Städte, dass die Einkommenssteuern und der Anteil der Einwohner mit höherem Einkommen hoch korreliert sind.

Der Zusammenhang zwischen Einkommensverteilung und Steuerbelastung zeigt sich für Schweizer Metropolregionen bereits bei einer deskriptiven Betrachtung sehr deutlich. Abbildung 1.3 zeigt Gemeindecharakteristika in der Metropolregion Zürich.[13] Die Steuerbelastung (Steuerfüsse) ist in den Karten oben rechts abgebildet. Die Zentrumsgemeinde Zürich ist von einem Kranz von steuergünstigen Gemeinden umgeben. Insbesondere die Gemeinden entlang des Zürichsees sind steuerlich privilegiert durch bis zu 35% tiefere Steuersätze. Die räumliche Einkommensverteilung, dargestellt in der Karte unten links, zeigt ein praktisch perfekt komplementäres Bild: Die Bewohner von steuergünstigen Gemeinden weisen durchschnittlich wesentlich höhere Einkommen aus als diejenigen von Gemeinden mit hohen Steuern. Dies gilt nicht nur im Vergleich von Zentrum und Peripherie, sondern auch in den feinen Abstufungen der Peripheriegemeinden untereinander. Die Kapitalisierung von Steuerunterschieden in den Mietpreisen lässt sich in der Karte unten links ablesen. Die unterschiedlichen Wohnpreise zwischen den

[13] Die Daten der Karten in Abbildung 1.3 basieren auf folgenden Quellen: Pendler: Bundesamt für Statistik, Volkszählung 1990. Steuerbelastung: Statistisches Amt des Kantons Zürich, Steuerfüsse 1997. Einkommensverteilung: Eidgenössische Steuerverwaltung, Direkte Bundessteuer 1997/98 - Gemeinden. Mietpreise: Wüest und Partner, Zürich. Berücksichtigt sind alle Gemeinden, aus mindestens 1/3 der Arbeitnehmer in die Zentrumsgemeinde Zürich pendeln.

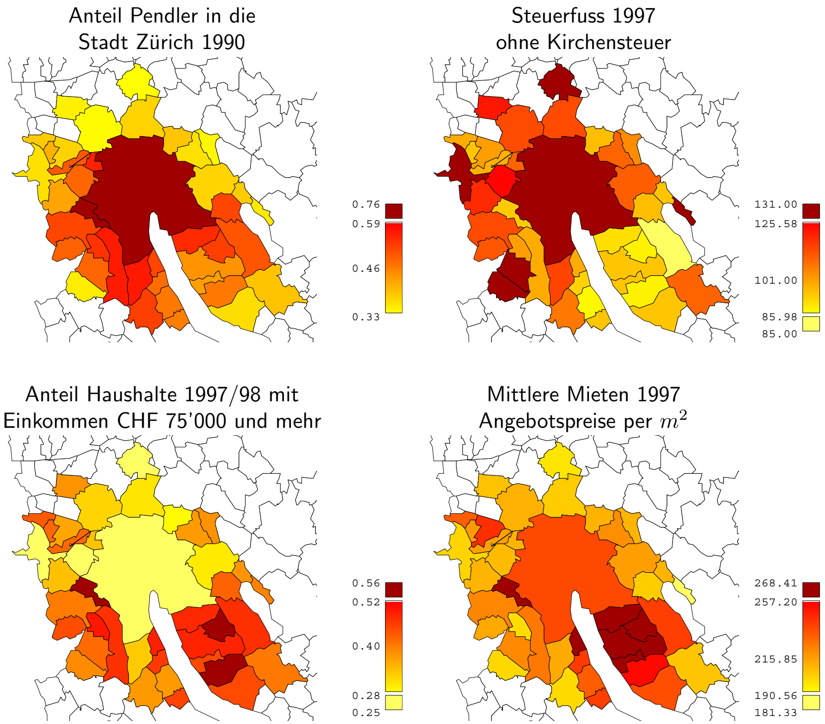

Abbildung 1.3: Gemeindecharakteristika in der Metropolregion Zürich.[13]

Gemeinden der Peripherie können gut durch Unterschiede in den Steuern erklärt werden. Der relativ hohe Preis für Wohnraum in der Zentrumsgemeinde trotz hohen Steuern lässt sich durch die allgemeine Attraktivität des Zentrums erklären.[14] Ein ähnliches Bild ergibt sich auch für die Metropolregion Basel in Abbildung 1.4.[15]

[14]Empirische Untersuchungen der Kapitalisierung, wie im Abschnitt oben besprochen, kontrollieren denn auch für alle wichtigen messbaren Standortfaktoren.

[15]Die Daten der Karten in Abbildung 1.4 basieren auf folgenden Quellen: Pendler: Bundesamt für Statistik, Volkszählung 1990. Steuerbelastung: Eidgenössische Steuerverwaltung, Steuerbelastung in der Schweiz, Natürliche Personen nach Gemeinden 1997 und eigene Berechnungen aufgrund der Gemeindesteuerfüsse der Kantone Basel-Stadt, Basel-Land und Solothurn. Einkommensverteilung: Eidgenössische Steuerverwaltung, Direkte Bundessteuer 1993/94 - Gemeinden. Mietpreise: Wüest und Partner, Zürich.

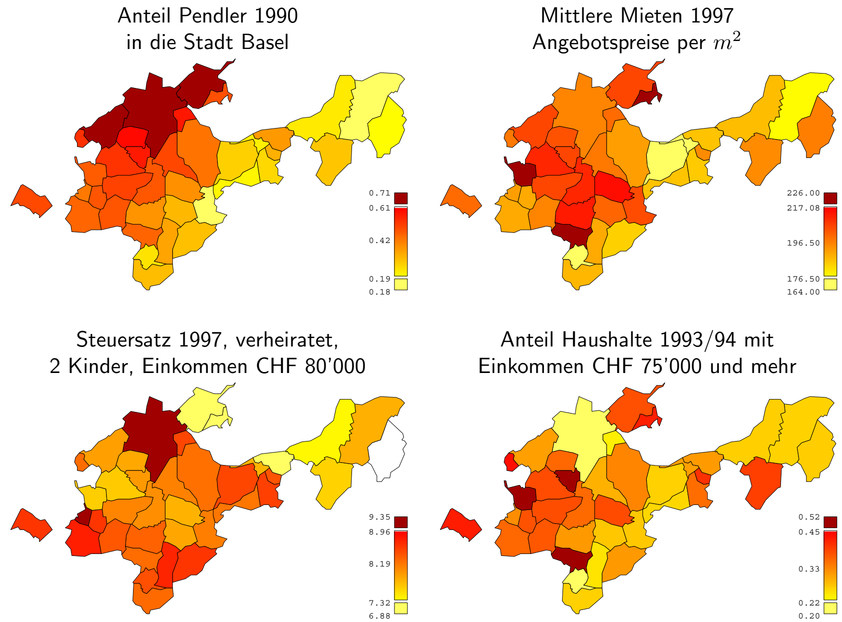

Abbildung 1.4: Gemeindecharakteristika in der Metropolregion Basel.[15]

Der für die beiden Metropolregionen deutlich sichtbare Zusammenhang zwischen Steuerbelastung und Einkommensverteilung könnte unter Umständen auch eine Folge von historisch bedingten Einkommensunterschieden sein. Die Steuerunterschiede ergäben sich dann dadurch, dass reiche Gemeinden aufgrund ihres 'guten' Steuersubstrats das gleiche Steueraufkommen mit tieferen Steuersätzen generieren könnten. Zur Beurteilung der theoretischen Modelle ist es deshalb wichtig zu überprüfen, ob die unterschiedliche Steuerbelastung nicht nur Folge, sondern auch Ursache der Einkommenssegregation ist.

In Kapitel 4 wird dazu die Wohnsitzwahl von individuellen Haushalten betrachtet, die im Jahr 1997 aus der Stadt Basel weggezogen oder innerhalb der Stadt umgezogen sind. Bereits die deskriptive Analyse liefert ein deut-

Berücksichtigt sind alle Gemeinden der Agglomeration Basel gemäss Definition des Bundesamtes für Statistik, basierend auf der Volkszählung 1990.

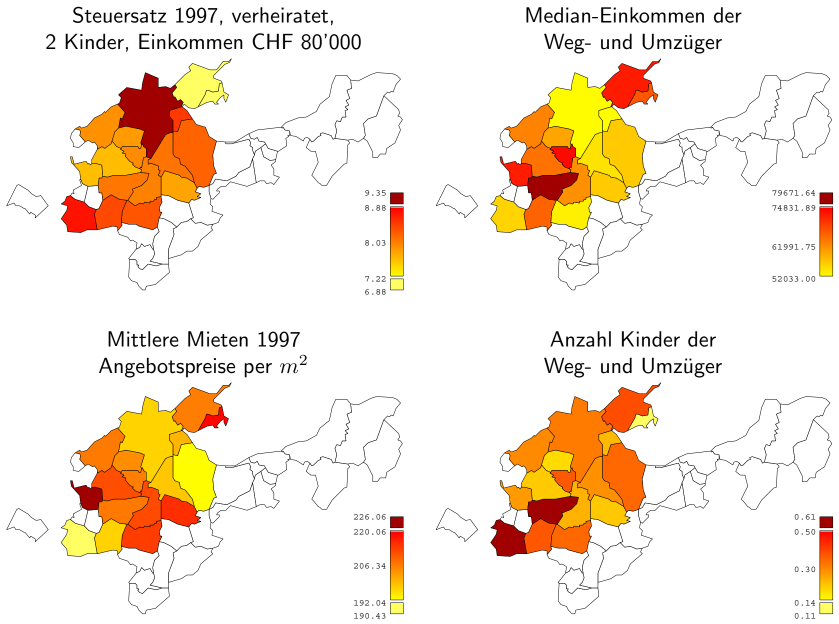

Abbildung 1.5: Weg- und Umzüger der Stadt Basel 1997.[16]

liches Bild. Abbildung 1.5 (obere Zeile) zeigt das mittlere Einkommen der Weg- und Umzüger im Vergleich mit den Steuersätzen.[16] Die beiden Karten machen deutlich, dass reiche Haushalte mit grösserer Wahrscheinlichkeit in eine steuergünstige Gemeinde der Peripherie ziehen als ärmere Haushalte. Dieser Zusammenhang wird besonders augenfällig im Vergleich von Zentrum und Peripherie. Die statistische Analyse in Kapitel 4 zeigt auf, dass dieser Zusammenhang hoch signifikant ist. Die unterschiedliche Familienstruktur der Haushalte (Abbildung 1.5, unten) deutet auf Präferenzsegregation hin, wie sie in Satz 5 postuliert wird.

[16]Die Daten der Karten in Abbildung 1.5 basieren auf folgenden Quellen: Steuerbelastung: Eidgenössische Steuerverwaltung, Steuerbelastung in der Schweiz, Natürliche Personen nach Gemeinden 1997. Einkommensverteilung und Anzahl Kinder: Eigene Berechnungen aufgrund von Daten der Steuerverwaltung und des Statistischen Amtes des Kantons Basel-Stadt. Mietpreise: Wüest und Partner, Zürich. Berücksichtigt sind alle (grösseren) Gemeinden, aus denen mindestens 36% der Arbeitnehmer in die Zentrumsgemeinde Basel pendeln.

1.6 Schlussbemerkungen

Städtische Agglomerationen der Schweiz bestehen aus einer Vielzahl von Gemeinden mit grossen Differenzen in der Steuerbelastung, in den Wohnpreisen und im durchschnittlichen Einkommen der Bewohner. Die Einwohner von steuergünstigen Gemeinden verfügen durchschnittlich über weit höhere Einkommen als die Einwohner von Gemeinden mit hohen Steuern.

Der theoretische Teil dieses Kapitels beschreibt die wechselseitige Abhängigkeit zwischen der räumlichen Einkommensverteilung und der lokalen Steuerbelastung in einem föderalistischen Gemeindesystem. Es wird aufgezeigt, dass sich ein Gleichgewicht mit lokalen Unterschieden in den Steuersätzen und im mittleren Einkommen ergibt, in dem kein Haushalt den gewählten Wohnort wechseln möchte. Anhand zweier theoretische Modelle wird erklärt, warum arme Haushalte im Gegensatz zu reichen Haushaltem nicht in die steuergünstigen Gemeinden ziehen. Diese endogene Einkommenssegregation entsteht im ersten Modell durch die Progressivität von lokalen Einkommenssteuern, im zweiten Modell durch die abnehmenden Wohnausgaben im Haushaltsbudget mit zunehmendem Einkommen. Die numerischen Simulationen in den kalibrierten Modellen zeigen, dass die in der Schweiz beobachteten Steuer- und Einkommensunterschiede durch die vorgestellten Modelle gut erklärt werden können.

Die empirische Untersuchung der Wohnsitzwahl von Haushalten in der Metropolregion Basel bestätigt, dass die tiefen Steuern in den 'reichen' Gemeinden nicht nur eine Folge des guten Steuersubstrats sind, sondern auch eine Ursache dieser räumlichen Einkommensunterschiede: Reiche Haushalte ziehen mit signifikant höherer Wahrscheinlichkeit in eine steuergünstige Gemeinde als ärmere Haushalte.

1.A Anhang

Formale Herleitung zu Abschnitt 1.2.2

Das Haushaltsproblem ist:

$$\max_{h,b} U(h,b,g;\alpha) = h^\alpha b^\alpha g^\gamma \quad N.B. \quad ph + b \leq y(1-t)\,.$$

Dies führt zur Nachfragefunktion nach Wohnraum

$$h^* = h(t,p;y\alpha) = \frac{\alpha y(1-t)}{p}$$

und zur indirekten Nutzenfunktion

$$V = \alpha^\alpha (1-\alpha)^{(1-\alpha)} p^\alpha y(1-t) g^\gamma\,.$$

Die marginalen Grenzraten der Substitution für Gemeindecharakteristika

$$\begin{aligned} M_{g,t} &:= \left.\frac{dg}{dt}\right|_{dV=0,dp=0} = -\frac{\partial V/\partial t}{\partial V/\partial g} = \frac{yg}{\gamma y(1-t)} = \frac{g}{\gamma(1-t)}\,, \\ M_{g,p} &:= \left.\frac{dg}{dp}\right|_{dV=0,dt=0} = -\frac{\partial V/\partial p}{\partial V/\partial g} = \frac{h^* g}{\gamma y(1-t)} = \frac{\alpha g}{\gamma p}\,, \\ M_{t,p} &:= \left.\frac{dt}{dp}\right|_{dV=0,dg=0} = -\frac{\partial V/\partial p}{\partial V/\partial t} = -\frac{h^*}{y} = -\frac{\alpha(1-t)}{p} \end{aligned}$$

sind unabhängig vom Einkommen y. Die Abhängigkeit dieser Grenzraten vom Einkommen ist jedoch eine notwendige Bedingung für Einkommenssegregation (vgl. Kapitel 2). Im Fall der Cobb-Douglas Nutzenfunktion kann dies direkt gezeigt werden. Dazu betrachten wir das Verhältnis der Nutzen $V_j(y) := V(p_j, t_j; y)$ und $V_i(y) := V(p_i, t_i; y)$, die ein Haushalt in den Gemeinden j und i erwartet:

$$dV_{ij}(y) := \frac{V_i(y)}{V_j(y)} = \frac{\alpha^\alpha(1-\alpha)^{(1-\alpha)} p_i^{-\alpha} y(1-t_i) g^\gamma}{\alpha^\alpha(1-\alpha)^{(1-\alpha)} p_j^{-\alpha} y(1-t_j) g^\gamma} = \left(\frac{p_i}{p_j}\right)^{-\alpha} \frac{1-t_i}{1-t_j}\,.$$

Da dieses Verhältnis unabhängig von y ist, bevorzugen alle Haushalte entweder die Gemeinde i ($dV > 1 \Rightarrow V_i > V_j$) oder die Gemeinde j ($dV < 1 \Rightarrow V_i < V_j$), oder alle Haushalte sind indifferent ($dV = 1 \Rightarrow V_i = V_j$).

Formale Beschreibung des Modells in Abschnitt 1.3

Dieser Anhang beschreibt das Modell mit progressiven Steuern formal für die Erweiterung mit heterogenen Präferenzen. Der einfache Fall mit homogenen Präferenzen ist ein Spezialfall davon.

Die Haushalte sind heteregon bezüglich des Einkommens $y \in [\underline{y}, \overline{y}]$, $0 < \underline{y}, \overline{y} < \infty$, und des Parameters $\alpha \in [0,1]$, der die Präferenz für Wohnen beschreibt. Einkommen und Präferenzen sind gemäss einer Dichtefunktion $f(y,\alpha) > 0$ verteilt. Im kalibrierten Modell wird die Verteilung von y und α als unabhängig angenommen: $f(y,\alpha) = f(y)\,f(\alpha)$.

Das Haushaltsproblem ist:

$$\max_{h,b} U(h,b,g;\alpha) = h^\alpha b^\alpha g^\gamma \quad N.B. \quad ph + b \leq y[1 - t \cdot r(y)]\,.$$

Der Steuertarif ist progressiv ausgestaltet: $\partial r(y)/\partial y > 0$. Der Durchschnittssteuersatz $t\,r(y)$ ist kleiner als der Grenzsteuersatz $t[y\,\partial r(y)/\partial y + r(y)]$ und beide liegen zwischen 0 und 1. Die Nachfragefunktion nach Wohnraum ist

$$h^* = h(t,p;y,\alpha) = \frac{\alpha y[1 - t\,r(y)]}{p}$$

und die indirekte Nutzenfunktion

$$V(p,t;y,\alpha) = \alpha^\alpha(1-\alpha)^{(1-\alpha)} p^{-\alpha} y[1 - t\,r(y)] g^\gamma\,.$$

Da in diesem Modell das Angebot an öffentlichen Gütern für alle Gemeinden identisch ist, betrachtet der Haushalt nur die marginale Grenzrate der Substitution zwischen dem lokalen Steuerfuss und dem lokalen Wohnpreis:

$$M_{t,p} := \left.\frac{dt}{dp}\right|_{dV=0,dg=0} = -\frac{\partial V/\partial p}{\partial V/\partial t} = -\frac{h^*}{y\,r(y)} = -\frac{\alpha[1 - t\,r(y)]}{p\,r(y)} < 0\,.$$

Das negative Vorzeichen bedeutet, dass höhere Steuern durch tiefere Wohnpreise kompensiert werden müssen, um Einwohner anzuziehen. Daraus folgt direkt Satz 2.

Die Grenzrate $M_{t,p}$ steigt mit zunehmendem Einkommen y

$$\frac{\partial M_{t,p}}{\partial y} = \frac{\partial r(y)}{\partial y} \frac{\alpha}{p\,r^2(y)} > 0$$

und fällt mit steigender Wohnpräferenz α

$$\frac{\partial M_{t,p}}{\partial \alpha} = -\frac{[1 - t\, r(y)]}{p\, r(y)} < 0\,.$$

Diese beiden Eigenschaften sind hinreichend und notwendig für die Existenz von Einkommens- und Präferenzsegregation, wie der allgemeine Beweis in Kapitel 2 zeigt. Im Fall der spezifizierten Cobb-Douglas Nutzenfunktion ist jedoch ein direkter Beweis für die Sätze 3 bis 5 möglich.

Beweis Satz 3 (Einkommenssegregation bei homogenen Präferenzen): Seien $p_j < p_i$ und $t_j > t_i$ die Wohnpreise resp. Steuerfüsse für ein beliebiges Paar von Gemeinden j und i sowie $V_j(y) := V(p_j, t_j; y)$ und $V_i(y) := V(p_i, t_i; y)$ die indirekten Nutzen eines Haushalts mit Einkommen y und Wohnpräferenz α in den zwei Gemeinden. Es lebe ein Haushalt mit Einkommen y' in Gemeinde j und ein Haushalt $y'' \neq y'$ in Gemeinde i. Es gilt deshalb für das Nutzenverhältnis der beiden Gemeinden

$$dV_{ij}(y) := \frac{V_i(y)}{V_j(y)} = \frac{\alpha^{\alpha}(1-\alpha)^{(1-\alpha)} p_i^{-\alpha} y [1 - t_i\, r(y)] g^{\gamma}}{\alpha^{\alpha}(1-\alpha)^{(1-\alpha)} p_j^{-\alpha} y [1 - t_j\, r(y)] g^{\gamma}} = \left(\frac{p_j}{p_i}\right)^{\alpha} \frac{1 - t_i\, r(y)}{1 - t_j\, r(y)}\,,$$

dass $dV(y'') > 1$ für Haushalt y'', der i bevorzugt und $dV(y') < 1$ für Haushalt y'. Da $dV_{ij}(y)$ eine stetige Funktion von y ist, muss es gemäss dem Zwischenwertsatz ein Einkommen $\hat{y} \in [y', y'']$ mit $dV(\hat{y}) = 1$ geben, d.h. es gibt einen *indifferenten Haushalt*, für den die Gemeinden gleichwertig sind. Die Ableitung des Nutzenverhältnisses nach y

$$\frac{\partial dV}{\partial y} = \left(\frac{p_i}{p_j}\right)^{-\alpha} \frac{(t_j - t_i)\, \partial r(y)/\partial y}{[1 - t_j\, r(y)]^2} > 0$$

wird durch den progressiv ausgestalteten Steuertarif positiv. Es gilt deshalb, dass alle (reichen) Haushalte mit einem Einkommen $y > \hat{y}$ und folglich $dV > 1$ die steuergünstige Gemeinde i bevorzugen und dass alle (armen) Haushalte mit $y < \hat{y}$ und folglich $dV > 1$ die preisgünstige Gemeinde j vorziehen. □

Beweis Satz 4 (Bedingte Einkommenssegregation): Für jede Subgruppe der Bevölkerung mit gleichen Präferenzen α kann der Beweis von Satz 3 angewendet werden. □

Beweis Satz 5 (Präferenzsegregation): Der Beweis ist analog zu Satz 3 unter Verwendung von

$$\frac{\partial dV}{\partial \alpha} = \left(\frac{p_i}{p_j}\right)^{-\alpha} \frac{1 - t_i\, r(y)}{1 - t_j\, r(y)} [ln(p_j) - ln(p_i)] < 0\,. \ \square$$

Die beiden Sätze 4 und 5 erlauben die Aggregierung der Bevölkerung n_j, der Steuereinnahmen T_j und der Wohnraumnachfrage HD_j in einer Gemeinde j:

$$n_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} f(y,\alpha)\, dy\, d\alpha\,,$$

$$T_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} t_j\, y\, f(y,\alpha)\, dy\, d\alpha = n_j\, t_j\, Ey_j\,,$$

$$HD_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} h(t_j, p_j, g_j; y, \alpha) f(y,\alpha)\, dy\, d\alpha\,,$$

wobei $\overline{y}_j(\alpha)$ und $\underline{y}_j(\alpha)$ den relativ reichsten (resp. ärmsten) Haushalt mit Wohnpräferenz α in der Gemeinde j bezeichnet. Dieser entspricht dem Haushalt, der bezüglich der nächststeuergünstigeren (resp. steuer*un*günstigeren) Gemeinde indifferent ist.

Die Produktionsfunktion des öffentlichen Gutes ist linear

$$C(g_j, n_j) = c_0 + c_1\, g_j\, n_j\,.$$

Das Angebot an Wohnraum steigt mit dem Preis

$$HS_j = L_j \cdot p_j^{\theta}\,.$$

Der Tarif der Einkommenssteuer wird hergeleitet in Young (1990) und erfüllt das Prinzip des 'equal sacrifice':

$$r(y) = r_0[1 - (1 + r_2\, y^{r_1})^{-1/r_1}]\,.$$

Der Durchschnittssteuersatz $t\, r(y)$ und der Grenzsteuersatz $t[y\, \partial r(y)/\partial y + r(y)]$ steigen stetig von 0 bis (asymptotisch) zu einem Maximum von $t\, r_0$, wobei der Grenzsteuersatz immer über dem Durchschnittssteuersatz liegt.

Kalibrierung des Modells in Abschnitt 1.3

Die Einkommensverteilung in der Bevölkerung lässt sich gut durch eine log-normale Verteilung beschreiben, die mit Hilfe von Daten der Schweizerischen Arbeitskräfteerhebung (SAKE 1995) parametrisiert wurde.[17] Das minimale Einkommen (y_{min}) von CHF 23'000 entspricht dem Existenzminimum eines Einpersonenhaushalts gemäss den Richtlinien der Schweizerischen Konferenz für Sozialhilfe SKOS, adjustiert für Inflation.

Der Parameter α in der Nutzenfunktion entspricht dem Budgetanteil für Wohnausgaben und beträgt gemäss SAKE für Haushalte in der Metropolregion Zürich im Mittel 0.25.[18] Eine Beta-Verteilung mit Mittelwert 0.25 und Standardabweichung 0.11 passt sich für die Modellerweiterung mit Präferenzheterogenität gut an die Daten an.

Die Parameter $r_0 = 0.135$, $r_1 = 1$ und $r2 = 0.00001$ beschreiben den Steuertarif des Kantons Zürich (Steuertarif Ledige gemäss Steuergesetz vom 8. Juni 1997, inkl. Änderung vom 19. November 2001) nahezu perfekt. Um im Referenzfall mit Steuerharmonisierung einen Steuerfuss von 1 zu generieren, wurde $r_0 = 0.132$ gesetzt. Der Parameter γ, der die Wertschätzung des öffentlichen Gutes beschreibt, und das Angebot an öffentlichen Gütern sind so festgelegt, dass im Gleichgewicht realistische Steuersätze entstehen. Die Preiselastizität des Wohnangebots ist $\theta = 3$. Die Produktionsfunktion des öffentlichen Gutes weist konstante Skalenerträge auf: $c_0 = 0, c_1 = 1$.

[17] Jahreseinkommen nach Bundessteuern und kantonalen Steuern. Die Verteilung wird unten bei CHF 23'000 und oben bei CHF 500'000 abgeschnitten. Dadurch wird die numerische Berechnung erleichtert und der Vergleich der Ergebnisse mit dem politökonomischen Modell in Abschnitt 1.4 ermöglicht, das ein existenzsicherndes Einkommen aller Haushalte voraussetzt.

[18] Ausgaben für Miete im Verhältnis zum Einkommen nach allen Steuern. 65% der Haushalte im SAKE Datensatz sind Mieter und stehen hier für die ganze Bevölkerung.

Chapter 2

Income Segregation in Multi-Community Models

2.1 Introduction

Over the past two decades, there has been ongoing research on multi-community models of urban agglomerations with heterogeneous agents. Multi-community models try to explain persistent differences in local tax levels and in the provision of local public goods in metropolitan systems of communities. A main finding of the literature is that these differences are accompanied by segregation of the population into groups with similar incomes and tastes. This chapter establishes the general conditions that induce such a self-sorting process of the population.

The study of multi-community models originates in Tiebout's (1956) seminal work. While Tiebout was seeking to determine the optimal number of jurisdictions, *positive theories* of community systems take the political landscape, i.e. the number of jurisdictions and their physical size, as given. This strand of literature builds on Westhoff (1977), who analyzed a general equilibrium model with heterogeneous agents and a fixed number of competing jurisdictions. The individual communities provide a local public good which is financed by a proportional local income tax. The residents of a community agree on the tax rate and on the amount of public goods in a majority vote that respects the community's budget balance. Households choose the community that offers them the best combination of public goods provision and tax rate.

The consideration of local housing markets by Ellickson (1971) and Rose-Ackerman (1979) was a major step towards more realistic models of jurisdictional systems. This extension was accompanied by a shift away from the study of models with income taxation to models with *property taxation*. The shift was coherent with the U.S. institutional reality, but also a way to circumvent the technical problems associated with the housing market in income tax models. Property tax models have subsequently been investigated by Epple, Filimon and Romer (1984, 1993) and Epple and Romer (1991). This strand of literature has been comprehensively reviewed by Ross and Yinger (1999). The consideration of further dimensions of household heterogeneity by Epple and Platt (1998) allowed an empirical examination as undertaken by Epple and Sieg (1999) and Epple, Romer and Sieg (2001). There are only very few studies of models with local *income taxation* and housing markets, rare examples are Goodspeed (1989) and Hansen and Kessler (2001a). Besides posing a theoretically interesting problem, local income taxation exists in metropolitan areas in Switzerland.

The *segregation hypothesis* is the central proposition in multi-community models with heterogeneous households. Endogenous segregation means that different people choose different communities when the communities differ in tax rates, housing prices and public goods provision. The population is then segregated as groups of similar household attributes tend to live at the same places. While the Tiebout model focuses on preference heterogeneity, Ellickson and Westhoff turned the attention to income as the main dimension of difference.

Westhoff (1977) established income segregation by assuming that the *relative preference* for the public good varies with income. Westhoff's relative preference assumption is equivalent to the Spence-Mirrlees, also called single crossing condition, of incentive theory and information economics. The mathematical analogy is maintained in property tax models with a housing market. However, the same strategy does not generally apply in income tax models with housing markets. In Westhoff's original model, households make their residence choice by comparing a two-dimensional set of community characteristics, namely the public good and the income tax rate. This two-dimensionality is maintained in the property tax model, where households are only concerned with the after-tax housing price and not with

the tax rate per se. Unfortunately, this reduction of dimensionality is only possible in very simplistic models with income taxation. Households have to make a choice which takes account of the three dimensions of community characteristics, namely the tax rates, the housing prices and the public goods provision. This chapter investigates the segregation conditions in this general setting.[1]

The first part of this chapter establishes a set of segregation conditions formulated for a broad class of models, covering property tax models as well as income tax models. The property tax models in the existing literature and the few proposed income tax models can be treated as special cases of this general setting. This framework allows to study a variety of new and less restrictive income tax models. The second part of this chapter presents a series of models and shows how these models satisfy the conditions described in the previous section. While in some models the proposed conditions are very naturally satisfied, one has to make specific assumptions of household preferences in others. The main finding is that income segregation in income tax models can only be established under very specific assumptions of household preferences. A further result is that income segregation cannot be ensured with nonlinear, i.e. progressive or regressive, tax schemes.

2.2 The Model

The model economy is divided into J distinct communities. The area is populated by a continuum of heterogeneous households which differ in income $y \in [\underline{y}, \overline{y}]$, $\underline{y} > 0$, $\overline{y} < \infty$. Income is distributed according to the density function $f(y) > 0$. There are three goods in the economy: private consumption b, housing h and a local publicly provided good g. The latter may be a pure public good, a publicly provided consumption good or a pure transfer. It is local in the sense that it is only consumed by the residents of a community.

A household can move costlessly and chooses the community that maxi-

[1]Note that the natural generalization of the single crossing condition to multi-dimensional problems used in information economics cannot be adapted for the segregation condition. Guesnerie and Laffont(1984) describe how the single Spence-Mirrlees condition is adapted to a multi-dimensional decision space in mechanism design problems.

mizes its utility as place of residence. Each community j can individually set the amount of the local public good g_j and the local tax rate $t_j \in [0,1]$. This decision is made in a majority vote by the residents who respect the budget balance in the community. At this point it is not specified whether the tax is based on property or income. Each community has a fixed amount of land L_j from which housing stock is produced. Households may be renters or house owners. The price for housing p_j in community j is determined in a competitive housing market. The private good is considered as the numeraire. A community j is fully characterised by the triple (t_j, p_j, g_j). The set of all possible community characteristics is given by $\Gamma = [0,1] \times \mathbb{R}^{++} \times \mathbb{R}^{+}$.

2.2.1 Indirect utility

In the general framework, a household is described by its indirect utility function. Section 2.3 shows how the indirect utility function is derived from different sets of utility functions and household budget constraints. The indirect utility function is a function on $\Gamma \times \mathbb{R}^{+}$ such that:

$$(t,p,g,y) \;\to\; V(t,p,g,y)\,.$$

The indirect utility function is assumed to be twice continuously differentiable in all its arguments everywhere on its domain.

The following three assumptions on the form of the indirect utility function are necessary conditions for segregation of the population in equilibrium.

Assumption 1 *For all* $(t,p,g,y) \in \Gamma \times \mathbb{R}^{+}$

$$V_t(t,p,g,y) := \left.\frac{\partial V}{\partial t}\right|_{t,p,g,y} < 0\,,$$

$$V_p(t,p,g,y) := \left.\frac{\partial V}{\partial p}\right|_{t,p,g,y} < 0\,,$$

$$V_g(t,p,g,y) := \left.\frac{\partial V}{\partial g}\right|_{t,p,g,y} > 0\,.$$

Assumption 1 is the standard assumption about the influence of prices, taxes and public goods on the household's well-being. Property 1 follows

directly from applying Assumption 1 to the total differential of the indirect utility function.

Property 1 (Relative preferences)
For all $(t,p,g,y) \in \Gamma \times \mathbb{R}^{+}$

$$M_{g,t}(t,p,g,y) := \left.\frac{dg}{dt}\right|_{dV=0,dp=0} = -\frac{V_t}{V_g} > 0\,,$$

$$M_{g,p}(t,p,g,y) := \left.\frac{dg}{dp}\right|_{dV=0,dt=0} = -\frac{V_p}{V_g} > 0\,,$$

$$M_{t,p}(t,p,g,y) := \left.\frac{dt}{dp}\right|_{dV=0,dg=0} = -\frac{V_p}{V_t} < 0\,.$$

Property 1 states that a household can be compensated for a tax increase either by more public good provision or by lower housing prices. It also states that a household is indifferent to higher housing prices if it is compensated by more public good provision. $M_{\cdot,\cdot}$ is called the marginal rate of substitution between two community characteristics. The marginal rates are well defined due to the strict inequalities in Assumption 1.

Assumption 2 (Constant sign of relative preferences)
At least one of the following three alternatives (a), (b) and (c) holds:
(a) *For all* $(t,p,g,y) \in \Gamma \times \mathbb{R}^{+}$ *either*

$$(CS^{+}_{g,t}) \quad \frac{\partial M_{g,t}}{\partial y} > 0 \quad \textit{or} \quad (CS^{-}_{g,t}) \quad \frac{\partial M_{g,t}}{\partial y} < 0\,.$$

(b) *For all* $(t,p,g,y) \in \Gamma \times \mathbb{R}^{+}$ *either*

$$(CS^{+}_{g,p}) \quad \frac{\partial M_{g,p}}{\partial y} > 0 \quad \textit{or} \quad (CS^{-}_{g,p}) \quad \frac{\partial M_{g,p}}{\partial y} < 0\,.$$

(c) *For all* $(t,p,g,y) \in \Gamma \times \mathbb{R}^{+}$ *either*

$$(CS^{+}_{t,p}) \quad \frac{\partial M_{t,p}}{\partial y} > 0 \quad \textit{or} \quad (CS^{-}_{t,p}) \quad \frac{\partial M_{t,p}}{\partial y} < 0\,.$$

Note that $M_{t,p} = -M_{g,p}/M_{g,t}$ by definition and hence

$$\frac{\partial M_{t,p}}{\partial y} = -\frac{\partial M_{g,p}/\partial y \cdot M_{g,t} - M_{g,p} \cdot \partial M_{g,t}/\partial y}{M_{g,t}^2} .$$

Assumption 2 states that the household's relative preferences for community characteristics change systematically with income. Alternative $CS_{g,t}^+$ implies that a rich household has to be compensated for a tax increase by strictly more public goods than a poor household. Alternative $CS_{g,p}^+$ supposes that the compensation for higher housing prices by public goods strictly increases with income. Alternative $CS_{t,p}^+$ means that the tax cut compensating for higher housing prices is strictly more substantial for a rich household than for a poor household. The alternatives $CS_{\cdot,\cdot}^-$ are interpreted analogously. Assumption 2 requires specific assumptions on the form of the utility function and/or the budget constraint. In Section 2.3 several examples in which Assumption 2 is naturally satisfied are presented.

Assumption 3 (Proportional shift of relative preferences)
One of the following two alternatives (a) and (b) holds:
(a) *For all* $(t, p, g, y) \in \Gamma \times \mathbb{R}^+$ *either*

$$\frac{\partial M_{g,t}}{\partial y} = 0 \quad or \quad \frac{\partial M_{g,p}}{\partial y} = 0 \quad or \quad \frac{\partial M_{t,p}}{\partial y} = 0 .$$

(b) *For all* $(t, p, g, y) \in \Gamma \times \mathbb{R}^+$ *both*

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} \quad \text{and} \quad \frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y}$$

with $M_{t,g} = 1/M_{g,t}$ *are independent of* y.

Assumption 3a is a very strong assumption. Assumption 3b is a weaker but very technical. Both alternatives are difficult to interpret intuitively and seem difficult to justify empirically. However, as will become clear in the next section, Assumption 3 is an indispensable condition for the occurrence of segregation. Assumption 2 and Assumption 3 are a three-dimensional generalization of the single-crossing condition.

2.2.2 Location Choice

The indirect utility function V yields the utility of a household with income y in a community j with income tax t_j, housing prices p_j and public good provision g_j. V implicitly describes the indifference surfaces in the (t, p, g) space.

A household with income y chooses the community which maximizes the household's utility. Hence, given a set of community characteristics, (t_j, p_j, g_j) for $j = 1, ..., J$, a household prefers community j over community i if and only if

$$V(t_j, p_j, g_j, y) \geq V(t_i, p_i, g_i, y) \quad \textit{for all } i .$$

The graphical representation of the indirect utility function is extensively used in the following proofs and therefore introduced in detail. Figure 2.1 displays the indirect utility function graphically. The left picture in Figure 2.1 shows an indifference surface in the 3-dimensional (t, p, g) space. Point 1 represents a community with income tax t_1, housing price p_1 and public good provision g_1. The indifference surface $\bar{V}_{1,y} := \{(t, p, g) : V(t, p, g, y) = V(t_1, p_1, g_1, y)\}$ covers all community characteristics that yield the same utility as community 1 for a household with income y. Community 2 with (t_2, p_2, g_2) is an example of such a community. All triples above the indifference curves are preferred to community 1. As stated in Property 1, the indifference is increasing in p and in t.

The right picture in Figure 2.1 is an illustration of the same indifference surface in the 2-dimensional (t, g) policy space for a given level of housing prices p. Suppose for the moment that housing prices are fixed at p_1. The solid curve covers the set $\bar{V}_{1,y,p_1} := \{(t, g) : V(t, p_1, g, y) = V(t_1, p_1, g_1, y)\}$ and corresponds to the solid subset of $\bar{V}_{1,y}$ in the left picture. The solid curve represents all community characteristics that yield the same utility as community 1 given p_1. Consider now a different housing price level p_2. All community characteristics considered to be indifferent to community 1 given the new price p_2 are on the dashed curve. Community 2 with (t_2, p_2, g_2) is an example of such a community. The dashed curve lies above the bold line when Property 1 holds. Note that the indifference surface depends on the household's income y.

In the following propositions, the allocation of the households across

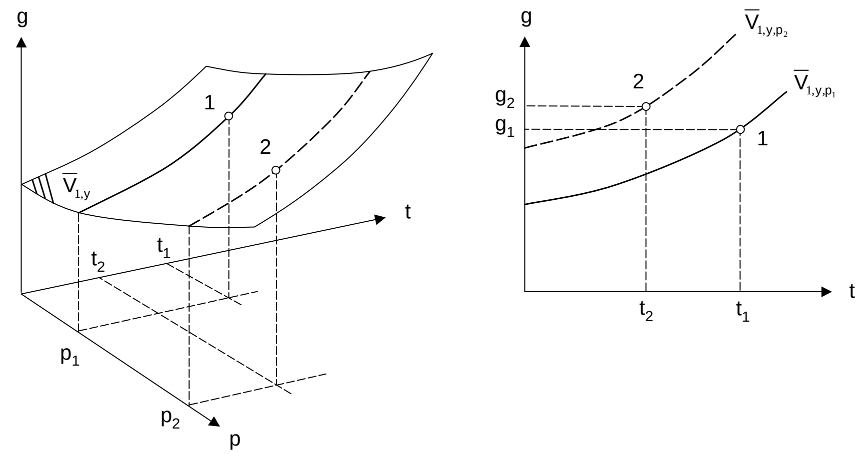

Figure 2.1: Indifference surface in the policy space.

distinct communities induced by the conditions in Assumptions 2 and 3 is discussed. A first observation is that all households are indifferent between all communities when the communities have identical community characteristics, i.e. $(t_j, p_j) = (t_i, p_i)$ for all j, i. In this case the households settle such that all communities show the same income distribution. This situation is a possible equilibrium in all the models presented in Section 2.3. In addition, it is always possible to think of equilibria in which subsets of communities have identical characteristics, i.e. $(t_j, p_j) = (t_i, p_i)$ for some j, i. However, these equilibria may not be stable.[2] The focus here is on the empirically interesting case of equilibria where all communities exhibit distinct characteristics.

[2]The notion of 'stability' in an intrinsically static model is rather peculiar. Nevertheless equilibria in static multi-community models are often judged by their 'dynamic' behavior. In this ad-hoc interpretation, an equilibrium is called 'stable' when the change of community characteristics induced by the migration of 'few' households gives these households an incentive to move back.

Proposition 1 (Boundary indifference)
When Assumption 1 holds and a household with income y' prefers to live in community j and another household with income $y'' > y'$ prefers to live in community i, then there exists a household with income $\hat{y}$, $y' \leq \hat{y} \leq y''$, which is indifferent between the two communities: $V(t_j, p_j, g_j, \hat{y}) = V(t_i, p_i, g_i, \hat{y})$.

Proof: Let $V_j(y) := V(t_j, p_j, g_j, y)$ be a household's utility in j and $V_i(y) := V(t_i, p_i, g_i, y)$ in i. The household with income y' prefers community j to i, hence $V_j(y') - V_i(y') \geq 0$. The opposite is true for a household with income y'': $V_j(y'') - V_i(y'') \leq 0$. $V_j(y) - V_i(y)$ is continuous in y since V is continuous in y. The intermediate value theorem implies that there is at least one $\hat{y}$ between y' and y'' s.t. $V_j(\hat{y}) - V_i(\hat{y}) = 0$. The existence of $\hat{y}$ follows from $f(y) > 0$. □

Proposition 1 states that for any pair of communities there is a 'border' household which is indifferent between the two.

Definition 1 (Perfect income segregation)
An allocation of households is called perfectly segregated by incomes if the J sets
$I_j = \{y :$ *household with income y prefers community j*$\}$ *satisfy*

- I_j *is an interval for all j,*
- $I_j \neq \emptyset$,
- $I_j \cap I_i = \emptyset$ *for all $j \neq i$,*
- $I_1 \cup ... \cup I_J = [\underline{y}, \overline{y}]$.

Definition 1 means that any community is populated by a single and distinct income class.

Proposition 2 (Perfect income segregation)
When Assumptions 1, 2 and 3 hold and all J communities exhibit distinct characteristics, $(t_j, p_j, g_j) \neq (t_i, p_i, g_i)$ for all $j \neq i$, then the allocation of households is perfectly segregated by incomes.

Proof: The proof proceeds in two steps. Firstly, income segregation is shown for a pair of two communities. Secondly, the result is extended to more than two communities.

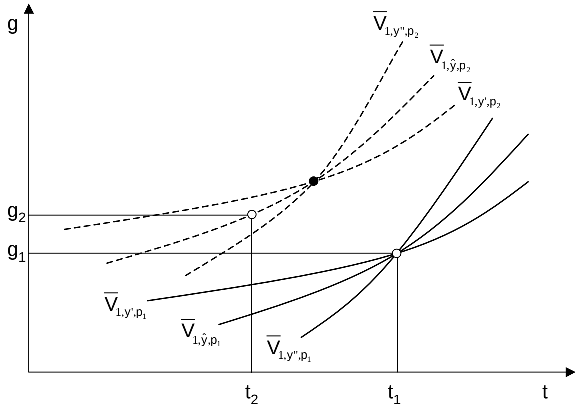

Figure 2.2: Indifference curves in the (t, g) policy space.

(1) The proof refers to Figure 2.2. Consider two communities 1 and 2 and assume $CS^+_{g,t}$. The figure shows the indifference surface in the (g, t) space for three different income levels $y' < \hat{y} < y''$ and two levels of housing prices $p_1 < p_2$. The solid lines represent all (t, g) pairs indifferent to community 1, characterized by (t_1, p_1, g_1), given p_1. The indifference curves are increasing in t (Property 1) and become steeper as income rises (assumption $CS^+_{g,t}$). The dashed lines represent all (t, g) pairs that are considered to be indifferent to community 1 given p_2. They are shifted to the left of the solid curves (Property 1) and intersect at the same point (Assumption 3 and proof in Appendix A). Imagine now a community 2, characterized by (t_2, p_2, g_2), which is considered as good as community 1 by household $\hat{y}$. If the (t_2, g_2) lies to the left of the intersection, then all richer households $y > \hat{y}$, e.g. y'', prefer community 2 to community 1 and all poorer households $y < y$, e.g. y', prefer community 1. If (t_2, g_2) is on the right side of the intersection the preference order is inverted. No segregation occurs in the unlikely case that (t_2, g_2) is exactly the intersection. The analogous argument holds for the other alternatives of Assumption 2.

(2) Suppose that the household allocation is not segregated: y' as well as y'' prefer community i, but y''' strictly prefers community j. Then it follows from Proposition 1 that there is a $\hat{y}$, $y' \leq \hat{y} < y'''$. (1) implies that

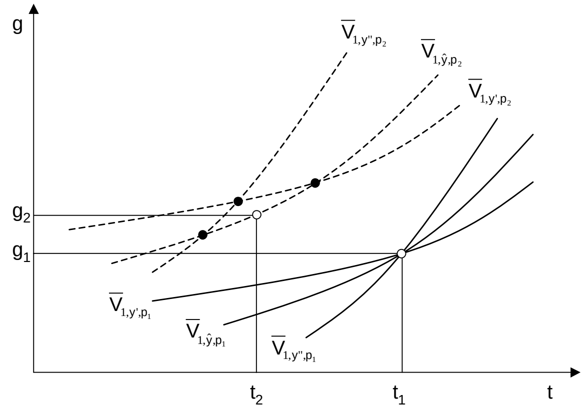

Figure 2.3: Indifference curves when Assumption 3 is not satisfied.

$y'' > y''' > \hat{y}$ strictly prefers j to i, which is a contradiction. □

Figure 2.3 shows a situation in which Assumption 3 does not hold. One can verify that the sketched indifference curves satisfy Assumption 2 ($CS^+_{g,t}$ and also $CS^+_{g,p}$ and $CS^+_{t,p}$). However, there is a richer household $y'' > \hat{y}$ that strictly prefers C_i and also a poorer household $y'' > \hat{y}$ that strictly prefers C_i. This contradicts Proposition 2 and shows that Assumption 3 is a necessary condition.

Note that very little can be said about the order of p, g and t across communities given the assumptions made so far. As Figure 2.2 shows, it is possible that rich people prefer the community with lower taxes or the one with higher taxes, that they live in the community with more or less public goods and that the rich communities have higher or lower housing prices. The properties of an equilibrium - if it exists - depends on the details of a fully specified model.

2.2.3 Taste Heterogeneity

In the previous sections, the sorting of the population with respect to income was discussed. There may be other sources of heterogeneity such as different tastes for housing or for the publicly provided good. Abstracting from income heterogeneity, one can use the above derived conditions and propositions for taste heterogeneity by simply replacing the variable y by a taste parameter, e.g. α, which satisfies all the assumptions respectively.

Combinations of income and taste heterogeneity result in more realistic models. Such models do not predict perfect segregation into income classes, but an allocation where the average household income still differs across communities. Models with both income and taste variation have been studied by Epple and Romer (1998), Epple and Sieg (1999), Epple, Romer and Sieg (2001), Kessler and Lülfesmann (1999) and in Chapter 3 of this book. These models form the basis for empirical investigation.

2.3 A Survey of Models

This section presents a series of models and discusses the specific assumptions under which these models lead to income segregation. The models are categorized along three dimensions: the nature of the local tax, the character of the publicly provided good and the form of the housing demand. The model properties are analyzed in in the three-dimensional (t, p, g) space of the general framework. Some models could also be examined in a two-dimensional characteristics space with the proofs becoming more elegant. However, it is the aim of this chapter to show the specifics of the whole model setup that generates segregation.

Each model is illustrated by an example. The household utility in these examples takes the form of a mixed CES/Stone-Geary function:

$$U(h, b, g) = [\gamma g^\rho + (1-\gamma) w(h, b)^\rho]^{1/\rho}$$

with

$$w(h, b) = (h - \beta_h)^\alpha (b - \beta_b)^{1-\alpha} .$$

Appendix B explains the characteristics of this utility function and derives the properties of the different models. There are two important features of the chosen utility function. Firstly, it allows the public good to be a perfect substitute $(\rho = 1, \sigma \to \infty)$, a perfect complement $(\rho \to -\infty, \sigma = 1)$ or of any intermediate degree of substitutability, which is measured by the elasticity of substitution $\sigma = 1/(1-\rho)$. Secondly, the budget share of

housing can change with income and therefore the income elasticity can differ from unity.

Table 2.1 summarizes the parameters of the specified models and shows their properties and predictions.

2.3.1 Property Tax, Transfer and Elastic Housing Demand

Epple and Romer (1991) propose a model in which the local jurisdictions pay a lump sum transfer g to their inhabitants.[3] The transfer is financed by a proportional property tax t on the consumption of housing. The individual households choose their place of residence and their optimal consumption bundle of a composite private good b and housing h. The household problem is

$$\max_{h,b} U(h, b) \quad s.t. \quad y + g \geq p(1+t)h + b,$$

where p is the pre-tax price of housing. The rate of substitution between the community variables p, g and t can now be derived under standard assumptions[4] by applying the envelope theorem to the total differential of the indirect utility:

$$M_{g,t} = ph^* , \quad M_{g,p} = (1+t)h^* , \quad M_{t,p} = -\frac{1+t}{p} ,$$

where $h^* = h(t, p, y)$ is the demand for housing. Assumptions $CS^+_{g,t}$ and $CS^+_{g,p}$ are satisfied as long as housing demand is strictly increasing in income. Assumption 3 holds as $M_{t,p}$ is independent of y.[5]

[3] Epple and Romer (1998) extend this model and allow for heterogeneous tastes. The utility function depends on a household specific parameter α, which describes the taste for housing:

$$\max_{h,b} U(h, b, \alpha) \quad s.t. \quad y + g \geq p(1+t)h + b .$$

Conditional income segregation arises for any given taste. Conditional taste segregation for a given income level is only established by explicitly making Assumption 2 in terms of α. Assumption 3 in terms of α is satisfied by construction.

[4] The standard assumptions are the following: the utility function is increasing, continuous and twice continuously differentiable in all its arguments.

[5] From a household's viewpoint, a community is fully characterized by the transfer and the after-tax housing price $p(1+t)$. Segregation could therefore also be analyzed in the two-dimensional $(g, p(1+t))$ space, which simplifies the proof of segregation.

From the point of view of a household a monetary transfer is equivalent to a perfectly substitutable public good. Hence one can study transfers in the specified example by assuming that $\sigma \to \infty$. The corresponding properties are reported as case [1] in Table 2.1. The example shows that the transfer g in the poor community, i.e. the community populated by households from lower income classes, is higher than in the rich community. The opposite can be stated for the after-tax price of housing $p(1+t)$, yet there is no resulting order for the tax rate t itself.

Other than in the original proof the general framework allows to study the effect of a nonlinear tax scheme. Consider that the property tax rate $t \cdot r$ consists of a tax shifter t, set by the community, and an exogenous tax rate structure $r(h)$ which can depend on the consumed amount of housing $h = h^*(y)$ and consequently on income. In this case, Assumption 3a is not satisfied any longer as $M_{t,p} = -[r^{-1}(h^*) + t]p^{-1}$ depends on y. Thus, income segregation cannot be established under a regressive or a progressive tax rate.

2.3.2 Property Tax, Public Good and Elastic Housing Demand

In the Epple, Filimon and Romer (1984, 1993) model, communities use the revenue from a proportional property tax to finance a local public good.[6] Differently from above, the public good enters the utility function and not the budget constraint:

$$\max_{h,b} U(h,b,g) \quad s.t. \quad y \geq p(1+t)h + b\,.$$

[6]Epple and Sieg (1999) extend this model and allow for heterogeneous tastes. The utility function depends on a household specific parameter α, which describes the taste for housing:

$$\max_{h,b} U(h,b,g,\alpha) \quad s.t. \quad y \geq p(1+t)h + b\,.$$

Conditional income segregation arises for any given taste. Conditional taste segregation for a given income level only occurs by explicitly making Assumption 2 in terms of α. Assumption 3 in terms of α is satisfied by construction.

The marginal rates of substitution between community characteristics become

$$M_{g,t} = p\,h^* \frac{U_b(h^*,b^*,g)}{U_g(h^*,b^*,g)}\,, \quad M_{g,p} = (1+t)h^* \frac{U_b(h^*,b^*,g)}{U_g(h^*,b^*,g)}\,, \quad M_{t,p} = -\frac{1+t}{p}\,.$$

Epple et al. make an explicit restriction on the household's preferences by assuming that $\partial[h^*\, U_b(h^*,b^*,g)/U_g(h^*,b^*,g)]/\partial y < 0$. This inequality guarantees that Assumptions $CS^-_{g,t}$ and $CS^-_{g,t}$ are satisfied. Assumption 3a is fulfilled by construction as $M_{t,p}$ is independent of income.[7]

The properties of different specifications of the example utility function are presented as cases [2] to [5] in Table 2.1. The example reveals that assumption made by Epple et al. is satisfied when the public good is either a complement (case [2]) or a substitute (case [4]). The order of community characteristics changes with the nature of the public good: Rich households will choose the community with high public good provision when they cannot easily substitute the public goods.

As in the previous section, income segregation cannot be established under a nonlinear tax scheme.

2.3.3 Income Tax, Transfer and Inelastic Housing Demand

Hansen and Kessler (2001a) present a model in which a pure monetary transfer is financed by a proportional local income tax. Hansen and Kessler assume that every household consumes one unit of housing independently of its income. This is a clearly unrealistic feature, but it allows a very elegant analysis.[8] The household problem is

$$\max_{b} U(h,b) \quad s.t. \quad y(1-t) + g \geq p + b\,.$$

The marginal rates of substitution between community characteristics are

$$M_{g,t} = y\,, \quad M_{g,p} = 1\,, \quad M_{t,p} = -1/y\,.$$

[7]Again, this model could be studied in the $(g, p(1+t))$ space.

[8]The housing price p in a community can formally be considered a reduction of the transfer g. This model can, therefore, easily be studied in the two-dimensional $(t, g-p)$ space.

Assumptions $CS^+_{g,t}$ and $CS^+_{g,p}$ are generically satisfied in this setup. Assumption 3a is fulfilled since $M_{g,p}$ is independent of y. Note, however, that the segregation conditions are only satisfied due to the extremely specific nature of the housing demand together with perfect substitutability of the public good.

Table 2.1 summarizes this model in case [6] and shows that poor households prefer the community with high transfer and high taxes.

2.3.4 Income Tax, Transfer and Elastic Housing Demand

The following model is a generalization of the model in the previous section. Housing demand is allowed to depend on income. The publicly provided good is still considered a pure transfer.[9] The household problem is

$$\max_{h,b} U(h,b) \quad s.t. \quad y(1-t) + g \geq ph + b\,.$$

The marginal rates of substitution between community characteristics are

$$M_{g,t} = y\,, \quad M_{g,p} = h^*\,, \quad M_{t,p} = -h^*/y\,.$$

Other than in the above models, this model cannot be analyzed in a reduced two-dimensional characteristics space. One can immediately see that assumption $CS^+_{g,t}$ is always satisfied and assumption $CS^+_{g,p}$ is met when housing demand is increasing in income. $CS_{t,p}$ is positive (negative) if the housing elasticity is smaller (bigger) than 1 since $\partial M_{t,p}/\partial y = h^*/y^2[1 - \partial h^*/\partial y \cdot y/h^*]$.

Assumption 3a is only fulfilled if housing demand h^* is a linear function of disposable income $y(1-t)$ through the origin, i.e. the preferences are homothetic and the income elasticity of housing is $\varepsilon_{h,y} = 1$ for all y. In this case $M_{t,p}$ is independent of y as h^*/y is a constant. Assumption 3b is satisfied if and only if housing demand is linear in income but not necessarily through the origin. Given this linearity $(\partial M_{g,p}/\partial y)/(\partial M_{g,t}/\partial y) = \partial h^*/\partial y$

[9] The same model is discussed in an unpublished study by Calabrese (1990). Unlike in the rest of the literature, Calabrese establishes the segregation conditions using restrictions from the public choice mechanism within the community. This proceeding may offer alternative ways to establish segregation conditions and deserves more attention.

and $(\partial M_{t,p}/\partial y)/(\partial M_{t,g}/\partial y) = y\partial h^*/\partial y - h^*$ are independent of y. Note that the assumption of linear housing demand is a necessary condition for segregation of the population in this model and not just a convenient simplification.

The example in Table 2.1 exemplifies the above reasoning in cases [7],[8] and [9]. The segregation conditions are always satisfied as the Stone-Geary subutility $w(h,b)$ generically leads to a linear housing demand function. If the income elasticity of housing is 1 (case [7]), then segregation is only driven by the nature of the public good. Hence, poor households prefer communities with higher transfers in this case.

Non-linear tax schemes are a main feature of a realistic income tax model. Consider that the income tax rate $t \cdot r$ consists of a tax shifter t, set by the community, and an exogenous tax rate structure $r(y)$ which depends on income. In this case, Assumptions 3a and 3b cannot be satisfied. Income segregation can therefore not be ensured under the prevalent progressive income tax schemes.

2.3.5 Income Tax, Public Good and Elastic Housing Demand

The attention is now turned to the most general model. Local jurisdictions provide a local public good financed by a proportional tax on income.[10] The household problem is the following:

$$\max_{h,b} U(h,b,g) \quad s.t. \quad y(1-t) \geq ph + b\,.$$

The marginal rates of substitution between community characteristics are

$$M_{g,t} = y\,\frac{U_b(h^*,b^*,g)}{U_g(h^*,b^*,g)}\,, \quad M_{g,p} = h^*\,\frac{U_b(h^*,b^*,g)}{U_g(h^*,b^*,g)}\,, \quad M_{t,p} = -\frac{h^*}{y}\,.$$

[10] This model has already been analyzed by Goodspeed(1989). The graphical proof of segration he provides in Goodspeed(1986) leads to Assumption 2 in this chapter. However, he fails to observe the importance of my Assumption 3. The Stone-Geary specification in the numerical simulation

$$\max_{h,b} U(h,b,g) = (h-\beta_h)^\alpha (b-\beta_b)^{1-\alpha} (g-\beta_g)^\gamma \quad s.t. \quad y(1-t) \geq ph + b$$

satisfies Assumption 3b by chance and thereby prevents the detection of the missing assumption.

There are several sets of assumptions that generate segregation in this model. The first assumption is to set the income elasticity of housing $\varepsilon_{h,y}$ to unity for all y. In this case, housing demand is a constant fraction of disposable income. As in Section 2.3.4 $M_{t,p}$ is independent of y and Assumption 3a is fulfilled. Assumption 2a and 2b are satisfied when $\partial[h^* \, U_b(h^*, b^*, g)/U_g(h^*, b^*, g)]/\partial y$ is either > 0 or < 0. With this set of assumptions the segregation of the population is fully driven by the nature of the public good. This situation and its qualitative implications for the equilibrium values are resumed in cases [10] to [12] in Table 2.1.

Another important source for segregation is the income elasticity of housing. Assume that the elasticity of substitution between g and $w(h, b)$ is exactly one. Table 2.1 shows the model properties if the income elasticity of housing is below unity in case[13] and if it is above unity in case [14]. Assumptions $CS^-_{g,t}$, $CS^-_{g,p}$, and $CS^+_{t,p}$ (case[13]), respectively $CS^-_{t,p}$ (case[14]) as well as Assumption 3b are satisfied without further assumptions.

If the public good is either a substitute or a complement (cases [15] and [16]), the segregation conditions are only satisfied if the income elasticity of housing is one. Although the utility function in case [15] satisfies $CS^-_{g,t}$, $CS^-_{g,p}$, and $CS^+_{t,p}$ it does fulfill neither Assumption 3a nor 3b. Therefore, given the specified utility function, segregation can be established if it is driven by either only the nature of the public good or by only the income elasticity of housing.

Note that segregation of the population cannot be established under a progressive or regressive income tax scheme as can easily be verified by looking at the conditions for Assumption 3 in Appendix B.

2.4 Conclusions

This chapter presents a general framework in which the segregation of the population induced by local tax setting in multi-community models with heterogeneous agents and housing markets can be analyzed. A set of conditions leading to segregation of the households is derived. While these conditions are naturally satisfied in property tax models, very restrictive assumptions on the household's preferences are needed in models with local income taxes.

The results in this chapter can be used to build more realistic multi-community models with local income taxation. Note that this paper does not investigate the entire general equilibrium model. Multi-community models have to be closed by specifying both the housing supply function and the public choice decision within communities. The latter is usually modelled as a majority vote. Existence of the equilibrium in the complete general equilibrium model is not generally guaranteed. Hansen and Kessler (2001b) show that segregation of the population is in many cases incompatible with the majority voting equilibrium within countries in models with local income taxation. Chapter 3 establishes equilibria in a model with local income taxes, a partly substitutable public good and elastic housing demand as outlined in Section 2.3.5.

The model class analyzed in this chapter is designed to explain decentralized public choice in metropolitan areas. For this purpose it is justifiable to consider the households' residence choice independent from the location of its members' jobs. However, the disregard of the dependence of residence choice from the availability of suitable jobs - and vice-versa - limits the usefulness of these models for examining fiscal decentralization on the level of federal states or countries. The consideration of the location choice of firms would be a major step towards a better understanding of fiscal decentralization on a national scale.

Table 2.1: Overview of model properties.

	linear property tax					linear income tax										
Section	2.3.1	2.3.2				2.3.3	2.3.4			2.3.5						
Case	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Restrictions on model parameters																
$\sigma_{w,g}$	∞	≥ 1	$= 1$	≤ 1	> 1	∞	∞	∞	∞	> 1	$= 1$	< 1	$= 1$	$= 1$	< 1	$\neq 1$
β_h	n.r.	≤ 0	$= 0$	≥ 0	> 0	> 0	$= 0$	> 0	> 0	$= 0$	$= 0$	$= 0$	> 0	> 0	> 0	$\neq 0$
β_b	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	$= 0$	> 0	> 0	$= 0$	$= 0$	$= 0$	> 0	> 0	> 0	$\neq 0$
ε_{h,y_d}	n.r.	n.r.	n.r.	n.r.	n.r.	$= 0$	$= 1$	< 1	> 1	$= 1$	$= 1$	$= 1$	< 1	> 1	< 1	$\neq 1$
Segregation conditions																
$\partial M_{g,t}/\partial y$	> 0	> 0	$= 0$	< 0	d.i.	> 0	> 0	> 0	> 0	> 0	$= 0$	< 0	< 0	< 0	< 0	d.i.
$\partial M_{g,p}/\partial y$	> 0	> 0	$= 0$	< 0	d.i.	$= 0$	> 0	> 0	> 0	> 0	$= 0$	< 0	< 0	< 0	< 0	d.i.
$\partial M_{t,p}/\partial y$	$= 0$	$= 0$	$= 0$	$= 0$	$= 0$	> 0	$= 0$	> 0	< 0	$= 0$	$= 0$	$= 0$	> 0	< 0	> 0	d.i.
Assumption 2	✓	✓	×	✓	×	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	×
Assumption 3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×
Ordering of community characteristics in equilibrium																
g^{poor} vs. g^{rich}	$>$	$>$	$=$	$<$	-	$>$	$>$	$\gtrless$	$\gtrless$	$>$	$=$	$<$	$\gtrless$	$\gtrless$	-	-
t^{poor} vs. t^{rich}	$\gtrless$	$\gtrless$	$=$	$\gtrless$	-	$>$	$\gtrless$	$\gtrless$	$\gtrless$	$\gtrless$	$=$	$\gtrless$	$\gtrless$	$\gtrless$	-	-
p^{poor} vs. p^{rich}	$\gtrless$	$\gtrless$	$=$	$\gtrless$	-	$\gtrless$	$\gtrless$	$\gtrless$	$\gtrless$	$\gtrless$	$=$	$\gtrless$	$\gtrless$	$\gtrless$	-	-
$p_{a.t.}^{poor}$ vs. $p_{a.t.}^{rich}$	$>$	$>$	$=$	$<$	-											

Notes: "n.r."indicates that no restrictions on the model parameter are made. "d.i."denotes that the sign depends on income, "✓"that an assumption is satisfied and "×"that an assumption is not satisfied. "$\gtrless$"means that the community characteristics differ but that their order is not determined by the restrictions made, "-"means that income segregation cannot be established. $g, t, p^{poor/rich}$ is the equilibrium public good provision, tax rate and housing price in the community with poorer/richer households. $p_{a.t.}$ stands for the net off property tax price of housing.

2.A Appendix A

This appendix shows the role of Assumption 3 in the proof of Proposition 2.

Consider three households with given income levels $y' < \hat{y} < y''$. The utility these households achieve in some community j with characteristics (t_j, p_j, g_j) is denoted by $V_j(y')$, $V_j(\hat{y})$ and $V_j(y'')$. The indifference surfaces associated with their achieved utilities in community 1 are implicitly defined as

$$\begin{aligned} f(t,p,g;y') &= V(t,p,g,y') - V_j(y') = 0\,, \\ f(t,p,g;\hat{y}) &= V(t,p,g,\hat{y}) - V_j(\hat{y}) = 0\,, \\ f(t,p,g;y'') &= V(t,p,g,y'') - V_j(y'') = 0\,. \end{aligned}$$

Note that $F : \Gamma \to \mathbb{R}^3$ defined by

$$F(t,p,g) = \begin{pmatrix} f(t,p,g;y') \\ f(t,p,g;\hat{y}) \\ f(t,p,g;y'') \end{pmatrix}$$

reaches $(0,0,0)'$ for $(t,p,g) = (t_j, p_j, g_j)$. Also note that F twice continuously differentiable in all its arguments since this was assumed for the indirect utility function. Given these definitions Assumption 3 can be stated in an alternative formulation.

Assumption 3 (Proportional shift, alternative formulation)

The Jacobian of F

$$DF(t,p,g) = \begin{bmatrix} V_t(t,p,g,y') & V_p(t,p,g,y') & V_g(t,p,g,y') \\ V_t(t,p,g,\hat{y}) & V_p(t,p,g,\hat{y}) & V_g(t,p,g,\hat{y}) \\ V_t(t,p,g,y'') & V_p(t,p,g,y'') & V_g(t,p,g,y'') \end{bmatrix}$$

is of rank 2 for any $(t,p,g) \in \Gamma$ and any triple of income levels $y' < \hat{y} < y'' \in \mathbb{R}^+$.

Assumption 3 states that - in a neighborhood around (t_j, p_j, g_j) and independent of the income levels - all three indifference surfaces *intersect in a common one-dimensional curve.* Note that they would intersect in single point in the case of full rank. As Assumption 3 holds for any (t, p, g) and any income levels the above argument extends to the whole interior of Γ.

The link of the above assumption to the formulation in Section 2.2.1 is discussed in the following. Note that the Jacobian cannot have rank 1 as the indifference surfaces do not coincide due to Assumption 2. Hence $Det(DF) = 0$ is a sufficient and necessary condition for Assumption 3 to hold.

Firstly, $Det(DF) = 0$ if two columns are proportional, i.e. either $V_t(t,p,g,y)/V_p(t,p,g,y)$, $V_t(t,p,g,y)/V_g(t,p,g,y)$ or $V_g(t,p,g,y)/V_p(t,p,g,y)$ is independent of y. This is equivalent to the formulation of Assumption 3a in Section 2.2.1.

Secondly, $Det(DF) = 0$ if one column can be expressed as a linear combination of the other two. For this to hold, the equation system in $\lambda_1, \lambda_2 \in \mathbb{R}$

$$\begin{aligned} V_p(t,p,g,y') &= \lambda_1 V_g(t,p,g,y') + \lambda_2 V_t(t,p,g,y') \\ V_p(t,p,g,\hat{y}) &= \lambda_1 V_g(t,p,g,\hat{y}) + \lambda_2 V_t(t,p,g,\hat{y}) \\ V_p(t,p,g,y'') &= \lambda_1 V_g(t,p,g,y'') + \lambda_2 V_t(t,p,g,y'') \end{aligned}$$

must have a unique solution. Solving the first two rows yields:

$$\lambda_1 = \frac{V_p(\hat{y})/V_t(\hat{y}) - V_p(y')/V_t(y')}{V_g(\hat{y})/V_t(\hat{y}) - V_g(y')/V_t(y')} = \frac{M_{t,p}(\hat{y}) - M_{t,p}(y')}{M_{t,g}(\hat{y}) - M_{t,g}(y')} \cong \frac{\partial M_{t,p}/\partial y}{\partial M_{t,g}/\partial y},$$

$$\lambda_2 = \frac{V_p(\hat{y})/V_g(\hat{y}) - V_p(y')/V_g(y')}{V_t(\hat{y})/V_g(\hat{y}) - V_t(y')/V_g(y')} = \frac{M_{g,p}(\hat{y}) - M_{g,p}(y')}{M_{g,t}(\hat{y}) - M_{g,t}(y')} \cong \frac{\partial M_{g,p}/\partial y}{\partial M_{g,t}/\partial y},$$

λ_1 and λ_2 will also solve the third row if they are independent of the income levels $\hat{y}$ and y'. This is equivalent to the formulation of Assumption 3b in Section 2.2.1.

2.B Appendix B

This Appendix shows the calculations for the specified mixed CES/Stone-Geary model in Section 2.3. Case numbers [.] refer to Table 2.1.

The utility function of a household is given by

$$U(h,b,g) = [\gamma g^\rho + (1-\gamma) w(h,b)^\rho]^{1/\rho}$$

and

$$w(h,b) = (h-\beta_h)^\alpha (b-\beta_b)^{1-\alpha},$$

where $\alpha, \gamma \in [0,1]$ and $\rho \in [-\infty, 1]$. $w(h,b)$ is the subutility from housing and the private good g. This specification assumes that the preferences are weakly separable between the public good and the other two goods (h,b). The subutility of h and b is of the Stone-Geary form (see for a discussion e.g. Deaton and Muellbauer 1980). Although the parameters β_h and β_b can be interpreted as subsistence quantities they need not be positive. For $\beta_h = \beta_b = 0$ the Stone-Geary utility function reduces to a Cobb-Douglas function.

w can be interpreted as quantity index of non-public goods. The preference for the public good and the non-public goods takes the CES form. $\sigma = 1/(1-\rho)$ is the elasticity of substitution between public and non-public goods. g and w are called substitutes when $\sigma > 1$ and complements for $\sigma < 1$. The CES function contains the Cobb-Douglas function $(\rho = 0, \sigma = 1)$, perfect complementarity $(\rho = -\infty, \sigma = 0)$ and perfect substitutability $(\rho = 1, \sigma = \infty)$ as special cases.

Property Tax

In the case of a linear property tax, the housing demand is

$$h^*(t,p,y) = \frac{\alpha[y - p(1+t)\beta_h - \beta_b]}{p(1+t)} + \beta_h = \frac{\alpha(y - y_s)}{p(1+t)} + \beta_h,$$

where $y_s = p(1+t)\beta_h - \beta_b$ are the minimal expenditures to reach the subsistence level. Note that the housing demand is independent of g because of the weak separability.

The marginal rates of substitution between the community characteristics are

$$M_{g,t} = \frac{[\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}(1+t)^{-\alpha}(y-y_s)]^\rho g^{1-\rho}(1-\gamma)ph^*}{\gamma(y-y_s)},$$

$$M_{g,p} = \frac{[\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}(1+t)^{-\alpha}(y-y_s)]^\rho g^{1-\rho}(1-\gamma)(1+t)h^*}{\gamma(y-y_s)},$$

$$M_{t,p} = -\frac{1+t}{p},$$

where $h^* = h^*(t,p,y)$. Note that Property 1 is satisfied as long as the income is above subsistence level.

The segregation conditions of Assumption 2 are:

$$\frac{\partial M_{g,t}}{\partial y} = \frac{[\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}(1+t)^{-\alpha}(y-y_s)]^\rho g^{1-\rho}(1-\gamma)p(\rho h^*-\beta_h)}{\gamma(y-y_s)^2},$$

$$\frac{\partial M_{g,p}}{\partial y} = \frac{[\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}(1+t)^{-\alpha}(y-y_s)]^\rho g^{1-\rho}(1-\gamma)(1+t)(\rho h^*-\beta_h)}{\gamma(y-y_s)^2},$$

$$\frac{\partial M_{t,p}}{\partial y} = 0.$$

The following relationships hold if all households can afford the subsistence level, i.e $y > y_s$ and $h^* \geq \beta_h$:

$$\frac{\partial M_{g,t}}{\partial y}, \frac{\partial M_{g,p}}{\partial y} \begin{cases} >0 & \text{if} \quad \rho h^* > \beta_h \quad \text{cases [1], [2]} \\ =0 & \text{if} \quad \rho h^* = \beta_h \quad \text{case [3]} \\ <0 & \text{if} \quad \rho h^* < \beta_h \quad \text{case [4]} \end{cases}.$$

Note that Assumption 3a is naturally satisfied in the property tax model since

$$\frac{\partial M_{t,p}}{\partial y} = 0.$$

Income Tax

In the case of a linear income tax, the housing demand is

$$h^*(t,p,y) = \frac{\alpha[y(1-t) - p\beta_h - \beta_b]}{p} + \beta_h = \frac{\alpha[y(1-t)-y_s]}{p} + \beta_h,$$

where $y_s = p\beta_h - \beta_b$ are the minimal expenditures to reach the subsistence level. Note that the housing demand is independent of g because of the assumed weak separability.

The income elasticity of housing is

$$\varepsilon(t,p,y) = \frac{\partial h^*}{\partial y}\frac{y}{h^*} = \frac{\alpha y(1+t)}{\alpha[y(1+t) - p\beta_h - \beta_b] + p\beta_h}.$$

Assuming that all households can afford more than the minimum required quantities, e.g. $y(1-t) > y_s$ for all y, it follows

$$\varepsilon \lesseqqgtr 1 \quad \text{iif} \quad (1-\alpha)p\beta_h - \alpha\beta_b \gtreqqless 0.$$

The marginal rates of substitution between the community characteristics are

$$M_{g,t} = \frac{\{\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}[y(1-t)-y_s]\}^\rho g^{1-\rho}(1-\gamma)y}{\gamma[y(1-t)-y_s]},$$

$$M_{g,p} = \frac{\{\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}[y(1-t)-y_s]\}^\rho g^{1-\rho}(1-\gamma)h^*}{\gamma[y(1-t)-y_s]},$$

$$M_{t,p} = -\frac{h^*}{y}$$

and satisfy Property 1 as long as income is above subsistence level.

The segregation conditions of Assumption 2 are:

$$\frac{\partial M_{g,t}}{\partial y} = \frac{\{\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}[y(1-t)-y_s]\}^\rho g^{1-\rho}(1-\gamma)[\rho y(1-t)-y_s]}{\gamma[y(1-t)-y_s]^2},$$

$$\frac{\partial M_{g,p}}{\partial y} = \frac{\{\alpha^\alpha(1-\alpha)^{(1-\alpha)}p^{-\alpha}[y(1-t)-y_s]\}^\rho g^{1-\rho}(1-\gamma)(1-t)(\rho h^*-\beta_h)}{\gamma[y(1-t)-y_s]^2},$$

$$\frac{\partial M_{t,p}}{\partial y} = \frac{(1-\alpha)p\beta_h - \alpha\beta_b}{py^2}.$$

The following relationships hold if all households can afford the subsistence level, i.e $y(1-t) > y_s$ and $h^* \geq \beta_h$:

$$\frac{\partial M_{g,t}}{\partial y} \begin{cases} > 0 & \text{if} \quad \rho y(1-t) > p\beta_h + \beta_b \quad \text{cases [6]-[10]} \\ = 0 & \text{if} \quad \rho y(1-t) = p\beta_h + \beta_b \quad \text{case [11]} \\ < 0 & \text{if} \quad \rho y(1-t) < p\beta_h + \beta_b \quad \text{cases [12]-[15]} \end{cases},$$

$$\frac{\partial M_{g,p}}{\partial y} \begin{cases} > 0 & \text{if} \quad \rho h^* > \beta_h \quad \text{cases [7]-[10]} \\ = 0 & \text{if} \quad \rho h^* = \beta_h \quad \text{cases [6],[11]} \\ < 0 & \text{if} \quad \rho h^* < \beta_h \quad \text{cases [12]-[15]} \end{cases},$$

$$\frac{\partial M_{t,p}}{\partial y} \begin{cases} > 0 & \text{if} \quad \varepsilon < 1 \quad \text{cases [6],[8],[13],[15]} \\ = 0 & \text{if} \quad \varepsilon = 1 \quad \text{cases [7],[10]-[12]} \\ < 0 & \text{if} \quad \varepsilon > 1 \quad \text{cases [9],[14]} \end{cases}.$$

Case [6] is directly derived from section 2.3.3.

Neither Assumption 3a nor 3b is generally satisfied since

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} = \frac{(1-t)(\rho h^* - \beta_h)}{\rho y(1-t) - y_s}$$

and

$$\frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y} = \frac{\{\alpha^\alpha (1-\alpha)^{(1-\alpha)} p^{-\alpha} [y(1-t) - y_s]\}^\rho g^{1-\rho} (1-\gamma)[(1-\alpha)p\beta_h - \alpha\beta_b]}{\gamma p[\rho y(1-t) - y_s]}$$

depend on income.

Assumption 3a is satisfied if $\varepsilon = 1$ (cases [7], [10], [11], [12]) since

$$\frac{\partial M_{t,p}}{\partial y} = 0 ,$$

and if $\rho = 1$ and $\alpha = 0$ (thus $\sigma \to \infty$ and $\varepsilon = 0$, case [6]) since

$$\frac{\partial M_{g,p}}{\partial y} = 0 .$$

Assumption 3b is satisfied if $\rho = 0$ ($\sigma = 1$, cases [13], [14]) since

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} = \frac{(1-t)\beta_h}{p\beta_h + \beta_b} ,$$

$$\frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y} = \frac{[(1-\alpha)p\beta_h - \alpha\beta_b](1-\gamma)g}{p\gamma(p\beta_h + \beta_b)}$$

and if $\rho = 1$ ($\sigma \to \infty$, cases [8], [9]) since

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} = \frac{(1-t)\alpha}{p} ,$$

$$\frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y} = -\frac{[(1-\alpha)p\beta_h - \alpha\beta_b]\alpha^\alpha (1-\alpha)^{(1-\alpha)}}{p^{(1+\alpha)}} .$$

Chapter 3

Equilibrium, Segregation and Local Income Taxation When Households Differ in Both Preferences and Incomes

3.1 Introduction

Decentralized decision making at appropriate levels of government is in many countries viewed as an essential factor for good government. In some countries the federal system involves a decentralized fiscal structure. In the United States, for example, the central government raises progressive income taxes while the individual states collect retail sales taxes and many communities are financed by property taxes. Fiscal Federalism has recently been intensively debated in the European Union. On the one hand there are attempts to coordinate fiscal policies across EU member states. On the other, increased regional self-government, as implied by the subsidiarity principle, calls for some regional fiscal autonomy. Oates (1972) argues that local units deciding upon public programs are more likely to trade off costs against benefits if these programs are financed by local taxes.

The formal framework for the study of local provision of local public goods originates in Tiebout's (1956) seminal work. Tiebout showed that fiscal decentralization leads to an efficient provision of local public goods because people with similar preferences would settle in a particular location and vote for their desired level of public goods provision. Tiebout's result rests heavily on the assumption that households have equal incomes. When households differ in incomes, this result may no longer hold. In this situation it is advantageous for high-income households to locate with other high-income households and hence reduce their tax burden. Furthermore, low-income households may also prefer to locate with high-income households to benefit from their large tax base. So, while the rich try to gather in a rich community, the poor may seek to follow and hence leave an ever smaller and poorer community. However, as the literature outlined in the next paragraph has shown, this 'race to the bottom' need not take place. This chapter extends this finding to local income taxation.

Following Tiebout, there is a long tradition of modelling fiscal decentralization at community level. The consideration of heterogeneous household income by Ellickson (1971) and Westhoff (1977) moved the focus away from seeking optimal community size to the study of urban areas with given community borders. While this strand of literature was followed by a large number of studies investigating local *property taxation* (surveyed in Ross and Yinger 1999), there have been few contributions on local taxation of *income* (e.g. Hansen and Kessler 2001a).

Multi-community models with heterogeneous agents predict a segregation of the population by income, i.e. households of the same income group live in the same community. While differences of average income across local jurisdictions are typically observed, perfect segregation of income groups is not an empirical phenomenon. This incomplete income segregation of the population may be attributed to heterogeneous tastes for public goods and housing or to a certain preference for a particular place. Epple and Platt (1998) study a model with property taxation and show that the introduction of heterogeneous taste for housing indeed predicts a more realistic incomplete segregation of the population. Epple and Sieg (1999) test the predicted income segregation and show that it is able to explain the differences of income distributions across communities in the Boston metropolitan area.

This chapter follows Epple and Platt (1998) but introduces heterogeneous tastes in a multi-community model with local *income taxation* and a partly substitutable public good. The income taxation model has been investigated by Goodspeed (1989). This study generalizes Goodspeed's anal-

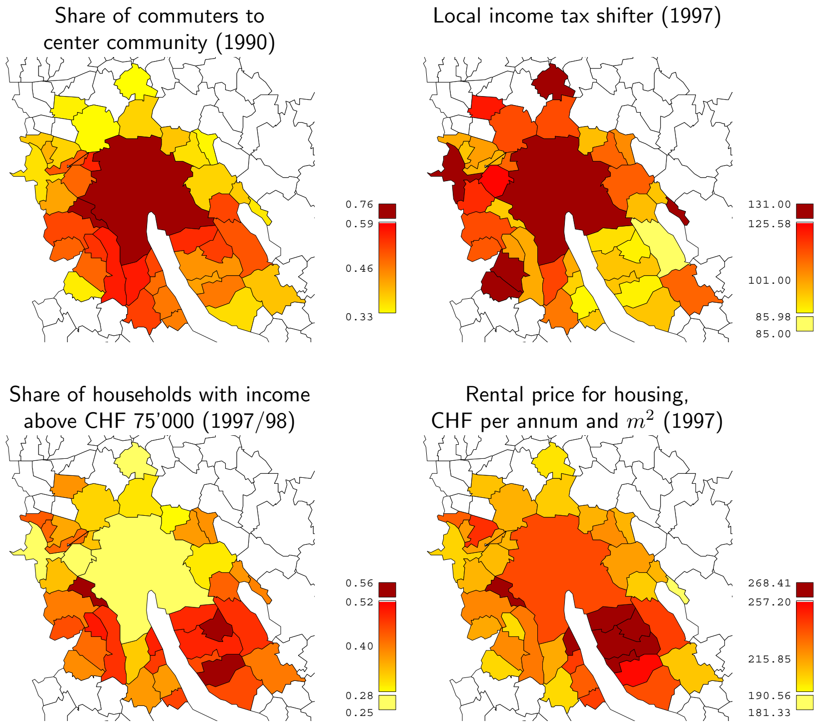

Figure 3.1: Community characteristics in the metropolitan area of Zurich.[2]

ysis both by introducing heterogeneous tastes and by using a realistic single-peaked distribution of the population. Not only does this single-peakedness capture a realistic feature of urban economies, but it also challenges the existence of equilibria in multi-community models with income taxation. The possible non-existence of segregated equilibria in a model with local income taxation is shown by Hansen and Kessler (2001b).

Switzerland is an exemplary case of a federal fiscal system. Switzerland is a federation of 26 states, the so-called cantons. The cantons are divided into individual communities of varying size and population. The roughly 3000 communities form individual jurisdictions with great autonomy in terms of providing local public goods such as school services or infrastructure. The unique situation in Switzerland is that the communities finance their expenditures mainly by local income taxes. While cantons autonomously organize the whole tax system, e.g. the degree of tax progression or the split between income and corporate taxes, the communities can generally only set a tax shifter in a given cantonal tax scheme. There is considerable variation in income taxes across Swiss communities. For example, for a two-child family with a gross income of 100,000 Swiss francs (CHF) the sum of cantonal and community income tax ranged from CHF 3,500 (city of Zug) to CHF 14,500 (city of Neuchâtel) in the year 2001.[1] Within metropolitan areas the (community) tax differences are smaller but may still differ by a factor of up to 1.5 in the Zurich area for example. Figure 3.1 shows the substantial differences in local tax levels, income and housing prices across this community system.[2] The bottom-left map visualizes the considerable segregation by incomes in the Zurich area. The top-right and the bottom-left maps demonstrate a striking relationship between income taxation and spatial income distribution: the local share of rich households is almost an inverted picture of the local tax levels. It is particularly interesting to see whether multi-community models are able to explain the observed tax differentials.

The chapter is organized as follows: Section 3.2 introduces the formal model and derives the properties of the household utility function that induce segregation of the population. In the first part of Section 3.3 the calibration used for the numerical investigation of the equilibrium is described. In the second part of Section 3.3 the numerical equilibrium is presented and the welfare implications of the decentralized decision making are discussed. Section 3.4 draws conclusions.

[1]Taken into consideration are the tax rates of the cantonal capitals. Some smaller communities show even higher respectively lower tax rates.

[2]Data from the following sources: Commuter: Swiss Federal Statistical Office, Census 1990. Tax rates: Statistisches Amt des Kantons Zürich, Steuerfüsse 1997. Income distribution: Swiss Federal Tax Administration. Housing prices: Wüest und Partner, Zurich. Considered are all communities where more than 1/3 of the working population is commuting to the center community.

3.2 The Model

The model economy is divided into J distinct communities. The area is populated by a continuum of heterogeneous households, which differ in both income $y \in [\underline{y}, \overline{y}]$, $0 < \underline{y}, \overline{y} < \infty$, and a parameter $\alpha \in [0, 1]$ describing their taste for housing. Income and taste are jointly distributed according to the density function $f(y, \alpha) > 0$. There are three goods in the economy: private consumption b, housing h and a local publicly provided good g.[3] The latter is local in the sense that it is only consumed by the residents of a community.

A household can move costlessly and chooses the community in which its utility is maximized as place of residence. Each community indexed by j can individually set the amount of the local public good g_j and the local income tax rate $t_j \in [0, 1]$. These decisions are made in a majority rule vote by the residents respecting budget balance in the community. Each community has a fixed amount of land L_j from which housing stock is produced. All households are renters and the housing stock is owned by an absentee landlord. The price for housing p_j in community j is determined in a competitive housing market. The private good is considered as numeraire. A community j is fully characterised by the triple (t_j, p_j, g_j). The set of all possible community characteristics is given by $\Gamma = [0, 1] \times \mathbb{R}^{++} \times \mathbb{R}^{+}$.

Location choice and voting are examined in a two-stage game. In the first stage, households choose their place of residence. In the second stage the inhabitants of a community vote for the level of public good provision and consequently for the community tax rate. The model is solved using backward induction.

3.2.1 Households

The preferences of the households are described by a Stone-Geary utility function

$$U(h, b, g; \alpha) := \alpha\, ln(h - \beta_h) + (1 - \alpha)\, ln(b - \beta_b) + \gamma\, ln(g - \beta_g)\,,$$

where h is the consumption of housing, b the consumption of the private good and g the consumption of the publicly provided good. $\beta_h > 0$, $\beta_b > 0$

[3]See Section 3.2.4 for a discussion of the nature of the public good.

and $\beta_g > 0$ are sometimes referred to as existential needs for housing, private good and public good, respectively. The parameter $\alpha \in [0,1]$ describes the households' taste for housing, as will become apparent below.

Households face a budget constraint:

$$ph + b \leq y(1-t)\,,$$

where p is the price of housing and t the local income tax. Note that the price of the private good is set to unity. Maximisation of the utility function with respect to h and b subject to the budget constraint yields the housing demand function

$$h^* := h(t,p,g;y,\alpha) = \frac{\alpha[y(1-t) - p\beta_h - \beta_b]}{p} + \beta_h$$

and the demand for the the private good $b^* = y(1-t) - ph^*$. Both demand functions are linear functions of after-tax income $y(1-t)$, reflecting the fact that a linear demand system implies a Stone-Geary utility function and vice-versa (see Deaton and Muellbauer 1980). Housing demand is increasing in α as long as the household can satisfy its existential needs, i.e. $\partial(h)/\partial(\alpha) > 0$ iff $y(1-t) > p\beta_h + \beta_b > 0$. $\alpha = 0$ implies that the housing demand is equal to the existential needs and hence does not change with household income. $\alpha = 1$ denotes a household which spends all extra income on housing after paying his existential need. The income elasticity of housing

$$\varepsilon := \frac{\partial h^*}{\partial y(1-t)} \frac{y(1-t)}{h^*} \lesseqgtr 1 \quad \text{iff} \quad \alpha \lesseqgtr \frac{p\beta_h}{p\beta_h + \beta_b}$$

depends on the housing price and is smaller or bigger than 1 depending on the household's tastes α (see the Appendix).

The indirect utility

$$V(t,p,g;y,\alpha) := U(h^*, b^*, g; \alpha)$$

gives the utility of a household with income y and preference parameter α in a community with income tax t, housing prices p and a public good provision g.

The following properties of the indirect utility function determine the distribution of households across communities. Properties 1 to 3 are directly derived from this indirect utility function assuming that existential needs are strictly satisfied, i.e. $y(1-t) > p\beta_h + \beta_b > 0$ and $g > \beta_g$. The calculations are provided in the Appendix.

Property 1 (Relative preferences)
For all $(t,p,g,y,\alpha) \in \Gamma \times \mathbb{R}^+ \times [0,1]$

$$M_{g,t}(t,p,g,y,\alpha) := \left.\frac{dg}{dt}\right|_{dV=0,dp=0} > 0\,,$$
$$M_{g,p}(t,p,g,y,\alpha) := \left.\frac{dg}{dp}\right|_{dV=0,dt=0} > 0\,,$$
$$M_{t,p}(t,p,g,y,\alpha) := \left.\frac{dt}{dp}\right|_{dV=0,dg=0} < 0\,.$$

Property 1 defines and signs the marginal rate of substitution $M_{.,.}$ between two community characteristics. Property 1 states that a household can be compensated for a tax increase either by more public good provision or by lower housing prices. Westhoff (1977) calls this trade-off the *relative preference* for the public good. Property 1 also states that a household can be made indifferent to higher housing prices if it is compensated by more public good provision. Property 1 holds under the standard assumption about the influence of prices, taxes and public goods on the household's well-being and is not specific to the assumed utility function.

Property 2 (Monotonicity of relative preferences)

(a) *For all* $(t,p,g,y) \in \Gamma \times \mathbb{R}^+$ *and any* $\alpha \in [0,1]$,

$$\frac{\partial M_{g,t}}{\partial y} < 0 \quad \text{and} \quad \frac{\partial M_{g,p}}{\partial y} < 0\,.$$

(b) *For all* $(t,p,g,\alpha) \in \Gamma \times [0,1]$ *and any* $y \in \mathbb{R}^+$,

$$\frac{\partial M_{g,p}}{\partial \alpha} > 0 \quad \text{and} \quad \frac{\partial M_{t,p}}{\partial \alpha} < 0\,.$$

Property 2 states that the relative preference for community characteristics changes *monotonically* with both income and taste. This property is equivalent to the Spence-Mirrless condition in information economics. It implies that a rich household can be compensated for a tax increase by strictly less public goods than a poor household. The compensation for higher housing prices by public goods decreases with income and increases with housing taste. The marginal rate of substitution between tax and housing price falls with housing taste.[4] Property 2 is a consequence of the changing weight of housing in the household's budget with income and taste. Note that Property 2 therefore does not hold for homothetic preferences.[5]

Property 3 (Proportional shift of relative preferences)

(a) *For all* $(t, p, g, y) \in \Gamma \times \mathbb{R}^+$ *and any given* $\alpha \in [0, 1]$*, both*

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} \quad \text{and} \quad \frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y}$$

are independent of y*, where* $M_{t,g} = 1/M_{g,t}$*.*

(b) *For all* $(t, p, g, \alpha) \in \Gamma \times [0, 1]$ *and any given* $y \in \mathbb{R}^+$*,*

$$\frac{\partial M_{g,t}}{\partial \alpha} = 0 \,.$$

Property 3 results from the linear demand system in combination with the additive separability between g and (h, b). Although very specific, Property 3 is an indispensable condition to get segregation of the population.[6]

3.2.2 Location Choice

A household chooses to locate in the community in which its utility is maximal. Defining $V_j(y, \alpha) := V(t_j, p_j, g_j; y, \alpha)$ as the household's utility in j, a

[4] The marginal rate of substitution between tax and housing price decreases with income if $\varepsilon > 1$ and increases if $\varepsilon < 1$.

[5] The property $\partial M_{g,t}/\partial y < 0$ is shared with Goodspeed (1986, 1989), who shows that it is equivalent to $\varepsilon_{g,y}/\varepsilon_{g,p} > 1$, where $\varepsilon_{g,y}$ is the income elasticity and ε_{g,p_g} is the (shadow) price elasticity of demand for the public good. $\partial M_{g,p}/\partial y < 0$ is shared with Ellikson (1971), Westhoff (1977), Epple, Filimon and Romer (1984, 1993) and Goodspeed (1986, 1989). Goodspeed reformulates this assumption as $\varepsilon_{g,y}/\varepsilon_{g,p} > \varepsilon_{h,y}$, where $\varepsilon_{h,y}$ is the income elasticity of demand for housing. He also points to empirical evidence that shows that both assumptions are reasonable. Property 2b shows preference heterogeneity in the same spirit as Epple and Platt (1998).

[6] It seems difficult to justify either Property 3a or Property 3b empirically. Chapter 2 shows that Property 3 is a necessairy condition for perfect income segregation. Goodspeed seems to derive income segregation without Property 3 in the same setting. However, the graphical proof he provides in Goodspeed (1986) is incomplete. Goodspeed (1989) uses the Stone-Geary utility function for the numerical simulation.

household chooses j if and only if

$$V_j(y, \alpha) \geq V_i(y, \alpha) \quad \text{for all } i \,. \tag{3.1}$$

The distribution of the households across communities implied by Properties 2 and 3 is described in the following paragraphs. A first observation is that all households are indifferent between all communities when the communities have identical community characteristics, i.e. $(t_i, p_i) = (t_j, p_j)$ for all j, i. In this case the households settle such that all communities show the same income distribution. In addition, it is always possible to think of equilibria in which subsets of communities have identical characteristics, i.e. $(t_i, p_i) = (t_j, p_j)$ for some j, i. However, these equilibria may not be stable.[7] The focus here is on the empirically interesting case of equilibria where all communities exhibit distinct characteristics.

The following paragraphs describe how the utility maximizing households will be allocated across communities.

Lemma 1 (Boundary indifference)
Consider the subpopulation with taste α*. If a household with income* y' *prefers to live in community* j *and another household with income* $y'' > y'$ *prefers to live in community* i*, then there is a 'border' household with income* $\hat{y}_{ji}(\alpha)$*,* $y' \leq \hat{y}_{ji}(\alpha) \leq y''$*, which is indifferent between the two communities.*

Proof: The household with income y' prefers j to i, hence $V_j(y') - V_i(y') \geq 0$. The opposite is true for a household with income y'' thus $V_j(y'') - V_i(y'') \leq 0$. $V_j(y) - V_i(y)$ is continuous in y as V is continuous in y. The intermediate value theorem implies that there is at least one $\hat{y}_{ji}$ between y' and y'' s.t. $V_j(\hat{y}_{ji}) - V_i(\hat{y}_{ji}) = 0$. The existence of $\hat{y}_{ji}$ follows from $f(y, \alpha) > 0$. □

The set of 'border' households is described by the function $\hat{y}_{ji}(\alpha)$. Equivalently, the set of border households is given by the inverse function $\hat{\alpha}_{ji}(y)$, implicitly defined by $V_j(\hat{\alpha}_{ji}(y)) = V_i(y, \hat{\alpha}_{ji}(y))$.

[7] The notion of 'stability' in an intrinsically static model is rather peculiar. Nevertheless, equilibria in static multi-community models are often judged by their 'dynamic' behavior. In this ad-hoc interpretation, an equilibrium is called 'stable' when the change of community characteristics induced by the migration of 'few' households gives these households an incentive to move back.

Definition 1 (Conditional income segregation)
An allocation of households is called conditionally segregated by incomes if the J sets $I_j = \{y :$ household with income y and taste α prefers community $j\}$ satisfy

- I_j *is an interval for all* j,
- $I_j \cap I_i = \emptyset$ *for all* $i \neq j$,
- $I_1 \cup ... \cup I_J = [\underline{y}, \overline{y}]$

for any α *and for any* j*:* $I_j \neq \emptyset$ *for at least one* α.

Definition 1 means that in a subpopulation with equal tastes any community is populated by a single and distinct income class.

Proposition 1 (Conditional income segregation)
When the household preferences are described by a Stone-Geary utility function and all J communities exhibit distinct characteristics, $(t_j, p_j, g_j) \neq (t_i, p_i, g_i)$ for all $i \neq j$, then the allocation of households is conditionally segregated by incomes.

Proof: The proof uses the fact that the utility difference $V_j - V_i = V(t_j, p_j, g_j; y, \alpha) - V(t_i, p_i, g_i; y, \alpha)$ between community j and i is strictly monotonic in y (see the Appendix):

$$sign\, \frac{\partial V_j - V_i}{\partial y} = sign(\frac{p_j \beta_h + \beta_b}{1 - t_j} - \frac{p_i \beta_h + \beta_b}{1 - t_i}).$$

Consider three households with income $y' < y'' < y'''$ respectively and suppose that the communities are not formed of non-overlapping intervals: y' as well as y''' prefer community j, but y'' strictly prefers community i. Given the opposed preference of y' and y''' it follows from Lemma 1 that there is an indifferent household $\hat{y}$, $y' \leq \hat{y} < y''$. The above sign condition implies that all households richer than $\hat{y}$, e.g. y''' , also prefer i, which is a contradiction. □

Chapter 2 shows that the Properties 1, 2a to 3a are sufficient conditions for income segregation.

Definition 2 (Conditional taste segregation)
An allocation of households is called conditionally segregated by tastes if the J sets $I_j = \{\alpha :$ household with income y and taste α prefers community $j\}$ satisfy

- I_j *is an interval for all* j,
- $I_j \cap I_j = \emptyset$, *for all* $i \neq j$,
- $I_1 \cup ... \cup I_J = [0, 1]$

for any y *and for any* j*:* $I_j \neq \emptyset$ *for at least one* y.

Definition 2 means that in a subpopulation with equal incomes any community is populated by a single and distinct interval of tastes.

Proposition 2 (Conditional taste segregation)
When the household preferences are described by a Stone-Geary utility function and all J communities exhibit distinct characteristics, $(t_j, p_j, g_j) \neq (t_i, p_i, g_i)$ for all $i \neq j$, then the allocation of households is conditionally segregated by tastes. Households in communities with lower housing prices have stronger tastes for housing than households in communities with higher prices.

Proof: The proof of the first sentence is analogous to Proposition 1 using the sign condition (derived in the appendix) $sign(\partial(dV_j - dV_i)/\partial\alpha) = sign(p_i - p_j)$. Second sentence: Consider $p_i < p_j$ and a household $(\hat{y}, \hat{\alpha})$ which is indifferent between the two communities j and i, hence $V_j(\hat{y}, \hat{\alpha}) = V_i(\hat{y}, \hat{\alpha})$. Then any household with the same income y and taste parameter $\alpha > \hat{\alpha}$ prefers community i, i.e. $V_j(\hat{y}, \hat{\alpha}) < V_i(\hat{y}, \hat{\alpha})$, since $\partial(dV_j - dV_i)/\partial\alpha < 0$ if $p_i < p_j$. □

Chapter 2 shows that the Properties 1, 2b to 3b are sufficient conditions for segregation by tastes.

Propositions 1 and 2 offer two ways of calculating a community's population:

$$n_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} f(y, \alpha)\, dy\, d\alpha = \int_{\underline{y}}^{\overline{y}} \int_{\underline{\alpha}_j(y)}^{\overline{\alpha}_j(y)} f(y, \alpha)\, d\alpha\, dy,$$

where $\underline{y}_j(\alpha)$ and $\overline{y}_j(\alpha)$ are the lowest and highest income in community j given the subpopulation with taste α. $\underline{y}_j(\alpha)$ is given by the locus of indifferent households $\hat{y}_{ji}$ between community j and its 'adjacent' community i

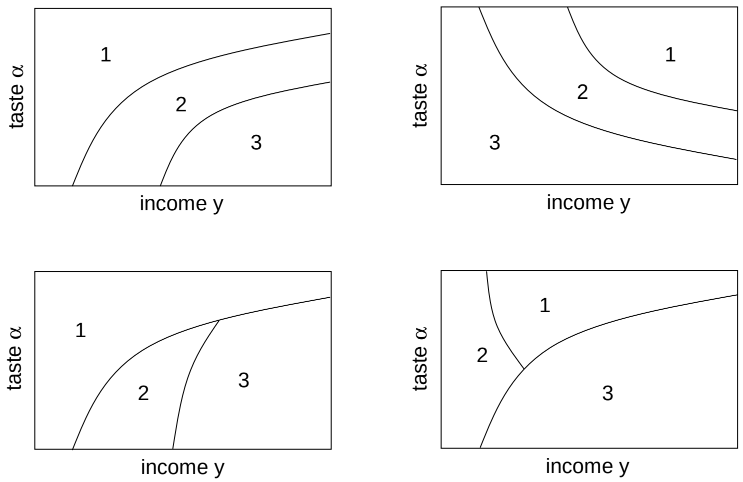

Figure 3.2: Examples of segregation patterns in the three-community case. The areas denoted by '1', '2' and '3' show the attributes of the households that prefer community 1, 2 or 3 respectively.

with lower income households. The other boundaries $\overline{y}_j(\alpha)$, $\underline{\alpha}_j(y)$ and $\overline{y}_j(\alpha)$ are given analogously. Note that the adjacent community might not be the same for all subpopulations. This is demonstrated in Figure 3.2 showing four examples of possible segregation patterns in the case of three populated communities.

3.2.3 Housing market

Within each community housing is produced from land and non-land factors. The housing supply in each community j is an increasing function of the housing price p_j and the land dedicated to housing L_j. The housing supply function

$$HS_j = L_j \cdot p_j^{\theta}$$

is adopted from Epple and Romer (1991), who derive the supply function from an explicit production function, where θ is the ratio of non-land to land input.

The aggregate housing demand in community j is

$$HD_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} h(t_j, p_j, g_j; y, \alpha) f(y, \alpha)\, dy\, d\alpha.$$

In equilibrium, the price for housing in community j clears the housing market

$$HD_j = HS_j. \qquad (3.2)$$

Definition 3 (Housing market tightness)
The housing market in community j with given population is called not too tight if

$$\frac{dp_j}{d(1-t_j)} \frac{1-t_j}{p_j}\Big|_{HD_j=HS_j} < 1\,.$$

Definition 3 defines the housing market as *not too tight* if the reaction of housing prices to changes in the tax rate and hence in the disposable income of the population is moderate. Note that the reaction of the housing price depends not only on the housing supply function but also on the characteristics, i.e. tastes and incomes, of the local population.

3.2.4 Public Choice

A community j provides a certain amount of a local public good to all its residents. The cost of providing this good is an increasing function of the amount provided g_j and the number of inhabitants n_j in the community. For simplicity, a linear function is assumed:

$$C(g_j, n_j) = c_0 + c_1\, g_j\, n_j\,,$$

where $c_0 \geq 0$ and $c_1 > 0$. Note that there are no spillovers in the production of the good across communities. The increasing cost in the number of beneficiaries means that the good is not a pure public good since there is rivalry in consumption. It is public in the sense that it is publicly provided and that all residents consume the same amount of the good. One can think of e.g. schools, street construction and maintenance, city planning activities, etc. A positive constant c_0 implies increasing returns to scale in the production of the public good.

The community finances the publicly provided good by a proportional income tax. The tax revenue is

$$T_j = \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} t_j\, y\, f(y,\alpha)\, dy\, d\alpha = n_j\, t_j\, Ey_j\,,$$

where Ey_j is the mean income in community j. In equilibrium, the community's budget is balanced:

$$C(g_j, n_j) = T_j\,. \tag{3.3}$$

The tax rate and the amount of public goods are determined in a majority rule vote by the residents of the community. At this stage, households take the population of the community as given.

Definition 4 (Public choice frontier PCF)
The public choice frontier PCF_j in community j is the set of (p_j, g_j, t_j) triples, where the pair (g_j, t_j) satisfies budget balance and p_j clears the housing market given the housing demand with tax rate t_j.

Proposition 3 (Segregation of voters)
Consider the subpopulation of households with taste α in community j and assume that the housing market is not too tight for all (p_j, t_j) on the PCF_j. If a household $\tilde{y}_j(\alpha)$ prefers the triple (p_j, g_j, t_j) on the PCF_j to all other triples on the PCF_j, then any richer (poorer) household opposes a reduction (increase) in taxes.

Proof: The proof refers to Figure 3.3. Consider the indifference curves of three voters with household income $y' < \tilde{y} < y''$ respectively, given the same taste parameter α. These indifference curves take into account the reaction of the housing prices to a change in the income tax rates. The straight line is the PCF in community j. One can verify in the figure that the pivotal voter $\tilde{y}$ prefers the pair (g_j, t_j) to all other combinations on the PCF. It is shown in the Appendix that the indifference curve is monotonically increasing in t and that its derivative w.r.t. t is decreasing in y. Therefore, all richer voters, e.g. y'', dislike all $(g, t) \in PCF_j$ combinations with taxes lower than t_j, while all poorer voters, e.g. y', dislike higher taxes. □

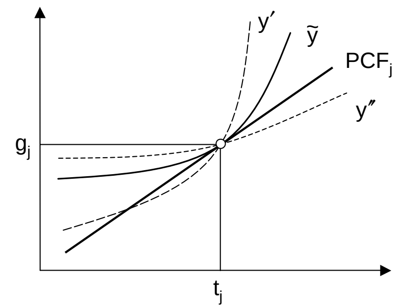

Figure 3.3: Voters' indifference curves in the (t, g) space.

$\tilde{y}_j(\alpha)$ is called the locus of pivotal voters. It is a decreasing function in α, as the price reduction induced by higher taxes is more appreciated by households with a stronger taste for housing. Note that from the perspective of a naïve voter who ignores the housing market, Proposition 3 holds without the additional assumption of the housing market tightness.

Definition 5 (Majority rule voting equilibrium)
A triple (p_j, g_j, t_j) on community j's PCF is called a majority rule voting equilibrium when no other triple on the PCF is strictly preferred by a majority of the community's residents.

As an implication of Proposition 3, a majority rule voting equilibrium in community j is established when

$$\int_0^1 \int_{\underline{y}_j(\alpha)}^{Min(\tilde{y}_j(\alpha), \overline{y}_j(\alpha))} f(y,\alpha)\, dy\, d\alpha = \frac{1}{2} \int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} f(y,\alpha)\, dy\, d\alpha \tag{3.4}$$

and the housing market is not too tight.

3.2.5 Equilibrium

The equilibrium as defined below is a subgame-perfect Nash equilibrium in the two-stage game of residential choice and voting.

Definition 6 (Equilibrium)
A set of community characteristics (p_j, g_j, t_j), $j = 1, ..., J$, *and an allocation of individual households across communities is an equilibrium if and only if*

- *all households choose their community to maximise their utility,*
- *the housing market clears in all communities,*
- *there is a majority rule voting equilibrium in all communities.*

Existence of the equilibrium is proofed by Goodspeed (1986) in a model with income taxes, taste homogeneity, naïve voters and a uniform income distribution. Epple, Filimon and Romer (1993) show existence in a model with property taxes and homogeneous tastes. Unfortunately, as in other models with taste heterogeneity (Epple and Platt, 1998), a proof of existence and uniqueness of this equilibrium could not be established. However, equations (1) to (4) provide the basis for a computational strategy to numerically find an equilibrium.

3.3 Numerical Equilibrium

In this section the qualitative and quantitative properties of the model are investigated in a fully specified model calibrated to the biggest Swiss metropolitan area.

3.3.1 Calibration

The area around the city of Zurich forms the biggest Swiss metropolitan area. The city of Zurich has about 330 thousand inhabitants and is the capital of the canton (state) of Zurich. The canton of Zurich counts 1.2 Million inhabitants in 171 individual communities. As described in the introduction, any of these communities can set an individual level of income taxes.

The analysis is restricted to the city of Zurich and a ring of the most integrated communities around the center. This ring is formed by all communities in the canton of Zurich with more than 1/3 of the working population commuting to the center.[8] Figure 3.4 shows a map with the city of Zurich

[8]The number of commuters to the city of Zurich and the size of the working population in the communities is based on the 1990 Census.

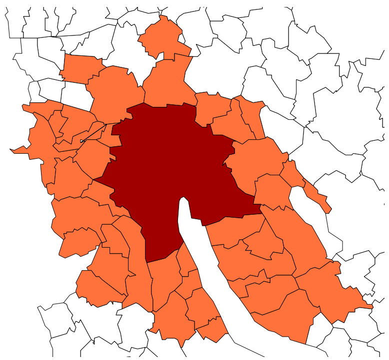

Figure 3.4: The Zurich metropolitan area around the lake of Zurich.

and the thus defined ring of 40 communities. The community characteristics of this area are also discussed in the introduction (see Figure 3.1).

The whole area has a physical size of $349km^2$, of which $88km^2$ (25%) form the city of Zurich. $140km^2$ are dedicated to development, $53km^2$ (38%) in the inner city and $87km^2$ in the fringe communities. In 1998, the whole area was populated by around 628'000 inhabitants, of whom 334'000 lived in the city and 294'000 in the fringe communities.[9] This agglomeration is modelled as two distinct jurisdictions with land size $L_1 = 0.4$ and $L_2 = 0.6$ respectively.

In the year 1997, the communal tax level in the fringe communities was on average 19% (minimum -35%, maximum +1%) lower than in the inner city.[10] The rental price for housing in the periphery was on average 6% lower than in the center (minimum -24%, maximum +13%).[11] The lowest-tax communities south-east of the city center exhibit substantially higher housing prices than the center. Figure 3.1 visualizes the spatial distribution of tax rates, incomes and housing prices in the area.

The income distribution is calibrated with data from the Swiss labor

[9]Source: Statistisches Amt des Kantons Zürich, Gemeindedaten per 31.12.1998.
[10]Source: Statistisches Amt des Kantons Zürich, Steuerfüsse 1997.
[11]Source: Wüest und Partner, Zurich. Offer prices for flats in newspapers and in the internet in 1997.

force survey.[12] The 1995 cross-section contains detailed information on 1124 households in the above defined region. These households had an average income of CHF 92,000 (median CHF 66,700) after state and federal taxes.[13] A log-normal distribution truncated at a minimum and a maximum income level is used to approximate the income distribution. A mean of $E(ln\, y) = 11.1$ and a standard deviation of $SD(ln\, y) = 0.55$ are close to the observed median and quartile distance. The minimum income is assumed to be $y_{min} = 23,000$ and the maximum income $y_{max} = 500,000$.[14] The median income in the city of Zurich is CHF 58,700 opposed to CHF 75,200 in the fringe communities.

The distribution of the taste parameter is described by a beta distribution. The Swiss labor force survey also contains the monthly housing expenditure of renters which allows to calibrate the distribution of tastes.[15] If the housing demand function is correctly specified, the taste parameter α of a household with disposable income y_d can be calculated as $(ph - ph_{min})/(y_d - y_{d,min})$, where ph is expenditure on housing and ph_{min} is the housing expenditure of the household with minimal disposable income $y_{d,min}$. The disposable income of a household y_d is calculated as reported household income minus federal, state and communal taxes. The average yearly housing expenditure of households around subsistence level is taken to approximate ph_{min}. This enables to approximate each household's taste parameter α. A beta distribution with mean $E(\alpha) = 0.17$ and standard deviation $SD(\alpha) = 0.11$ describes the distribution of the so calculated taste parameter fairly well. Taste and income are assumed to be uncorrelated.

The price elasticity of housing supply is $\theta = 3$ as in Epple and Romer (1991) and Goodspeed (1989). The production of the public good exhibits constant returns to scale, i.e. $c_0 = 0$ and $c_1 = 1$. $\beta_h = 700$, $\beta_b = 13000$ are chosen such that the consumption bundle of the minimal income household

[12]Swiss Federal Statistical Office, Schweizerische Arbeitskräfterhebung (SAKE) 1995.

[13]State and Federal taxes were deducted from net household income (after social security contribution) assuming a two-child family.

[14]The minimum income is subsistence level for a one-person-household as defined by the Schweizerische Konferenz für Sozialhilfe (SKOS) and adjusted for inflation. The maximum income is chosen arbitrarily, but has no influence on the numerical simulation due to the low weight on high incomes.

[15]Of course, there is a selection bias by only considering renters. Because the proportion of renters is very high in Switzerland (65% in the data set used), this is justified.

in equilibrium corresponds to the empirical findings. It is assumed that the existential needs for the public good are fairly high and that the benefit from additional units is limited: $\beta_g = 4000$ and $\gamma = 0.02$.[16] The assumed parameters result in equilibrium tax rates close to the observed ones. The parameters are summarized at the bottom of Table 3.1.

3.3.2 Simulation Results

The equilibrium values p_j, g_j and t_j, $i = 1, 2$, must satisfy equations (3.2), (3.3) and (3.4) and guarantee that the households reside in the community they prefer as expressed in equation (3.1). Unfortunately, there is no closed form solution to this nonlinear system of 6 equations and 6 unknowns. The equation system is therefore numerically solved for the equilibrium values of the model.[17]

Table 3.1 shows the equilibrium values for the calibrated model in columns 2 and 3. The equilibrium values for the case of a unified jurisdiction and for the case of taste homogeneity are given for comparison. As can be seen, the equilibrium values of the two communities differ substantially. I will refer to the high-tax community 1 as the 'center' and the low-tax community 2 as the 'periphery'.[18] The tax rate t_1 in the center is 44% higher than in the suburbs, whereas the housing price is 12% and the

[16]The properties of the equilibrium depends decisively on the preference parameters of the public good. The assumption that the existential need is small but the benefit from additional units is important can lead to numerical equilibria in which the high-tax communities exhibit higher housing prices but *lower* public goods provision. The rich households will then prefer the high-tax communities. This situation is in clear contradiction to the observed pattern in Swiss metropolitan areas.

[17]Numerically solving the equation system is tedious and time-consuming. The aggregation of individual demand and voting behavior requires double integrals over the community population. These integrals cannot be calculated analytically. Gauss-Legendre Quadrature with 40 nodes in each dimension is used to approximate the various double integrals. Numerically minimizing the sum of squared deviations from the equilibrium conditions with the Gauss-Newton method solves for the equilibrium values. Appropriate scaling of the arguments and of the equilibrium conditions is important for the accuracy of the result. Convergence is only achieved with 'good' starting values. Starting values are obtained from a grid search over the six-dimensional space of possible values. Different starting values lead to the same equilibrium values.

[18]These labels are arbitrary. There is always a second equilibrium with lower taxes in community 1. If the two communities have the same physical size, these two equilibria are identical except for the community index.

Table 3.1: Equilibrium values of the simulation.

		heterogen. tastes		homogen. tastes	
	unified	center	periphery	center	periphery
L: land size	1	0.40	0.60	0.40	0.60
p: housing price	11.40	10.48	11.92	9.19	12.37
t: income tax rate	0.064	0.085	0.059	0.110	0.056
g: public good prov.	5032	4488	5225	4335	5390
n: inhabitants	1	0.284	0.716	0.314	0.686
Ey: average income	78,547	52,995	88,687	39,368	96,460
Average CV		50.9	84.4	41.8	455.9
n (CV> 0)		0.171	0.506	0.180	0.686

The model parameters are: $\beta_h = 700$, $\beta_b = 13000$, $\beta_g = 4000$, $\gamma = 0.02$, $E(\alpha) = 0.17$, $SD(\alpha) = 0.11$ (heterogeneous tastes), $SD(\alpha) = 0$ (homogeneous tastes), $E(ln\, y) = 11.1$, $SD(ln\, y) = 0.55$, $y_{min} = 23'000$, $y_{max} = 500'000$, $\theta = 3$, $c_0 = 0$ and $c_1 = 1$.

public provision 14% lower. The average household income in the center is CHF 53 thousand a year compared to CHF 89 thousand in the suburbs. Thus, the simulated model is able to explain tax and income differences of the magnitude observed in the Zurich area. Not surprisingly, the model is very poor at explaining the high housing prices in the center as any immanent center advantages such as closeness to the main business and cultural activities are neglected.

The segregation of the population in the two communities is shown in the left picture in Figure 3.5. The locus of indifferent households, $\hat{y}_{12}$, turns out to be an increasing function of income in the present equilibrium. This implies that, given a subpopulation with equal tastes, richer households prefer the low-tax-high-price community.[19] However, this does not lead to a perfect income segregation between the two communities since the households have different preferences. Although the average income in the center is much lower than in the periphery, households from almost all income groups can be found in both communities. Figure 3.6 presents the resulting income distributions in the two communities. Figure 3.5 also shows

[19]Note that all households with a very high taste for housing prefer to live in the low-price community. This, however, applies to only 5% of the population, as the weight on taste parameters above 0.38 is low.

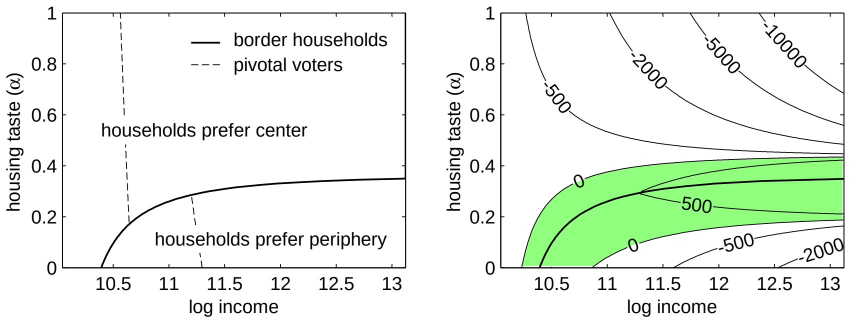

Figure 3.5: Income and taste segregation in equilibrium. The left figure includes the loci of pivotal voters. The right figure shows contour lines of the compensating variation (CV).

the loci of pivotal voters which split the communities' populations into half. Households in the 'rich' suburbs vote for more public goods than households in the 'poor' center, yet this generous public good provision can be financed by a lower tax rate, due to the higher average income of the residents.

The above segregated equilibrium is now compared to the equilibrium when the two jurisdictions were unified. The latter is equivalent to the equilibrium that would emerge if the two distinct jurisdictions harmonized their income tax levels and households located randomly. The equilibrium values of the unified community are presented in column 1 in Table 3.1. One can immediately see that the housing price, tax level and public good provision lie between the corresponding values in the two-community model. The competition of the two communities for households does thus not lead to an overall reduction of taxes, but to relatively lower taxes in the 'rich' community and higher taxes in the 'poor' community. The welfare effects associated with the segregated equilibrium depend on both the households' incomes and tastes.

The welfare effects are revealed by inspecting the compensating variation (CV) for the different types of households. The compensating variation is the additional gross income that a households needs in order to be compensated for a shift from the symmetric to the segregated equilibrium.[20] The

[20]The CV defined above is not a comprehensive measure of the welfare implications

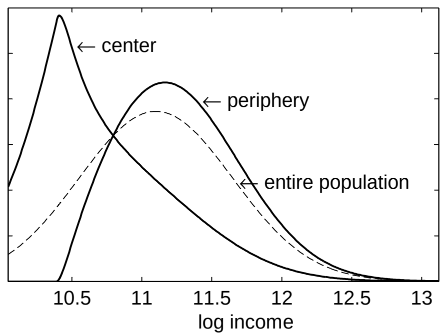

Figure 3.6: Income distribution in the center and the periphery.

right picture in Figure 3.5 shows contour lines of the CV. Households in the shaded band between the two zero contour lines exhibit positive values of the CV and thus prefer the one-community situation. Households further away from the border household prefer competing jurisdictions. The average CV's are reported in Table 3.1. Households in the 'poor' community have to be compensated by an average income allowance of CHF 51 compared to CHF 84 in the 'rich' community. Note that this amount is only one-tenth of a percent of average gross income. The number of households that prefer tax harmonization is also given in Table 3.1. 60% of the population in the 'poor' community and 71% of the population in the 'rich' community prefer tax harmonization.

How does the taste heterogeneity affects the properties of the equilibrium? The last two columns in Table 3.1 give the equilibrium values in the case of constant tastes that equal the average heterogeneous tastes. This equilibrium features perfect income segregation. Consequently, the income difference between the two communities is much larger than with heterogeneous tastes. Also, the price, tax and public good provision differences across the communities are stronger. The equilibrium values in a unified community equal the ones with taste homogeneity.[21] The welfare effects

since it ignores the welfare implications for the (absentee) landlord. The analytical solution for the CV given household characteristics is given in the Appendix.

[21]The symmetric equilibrium in the case of taste homogeneity is theoretically different

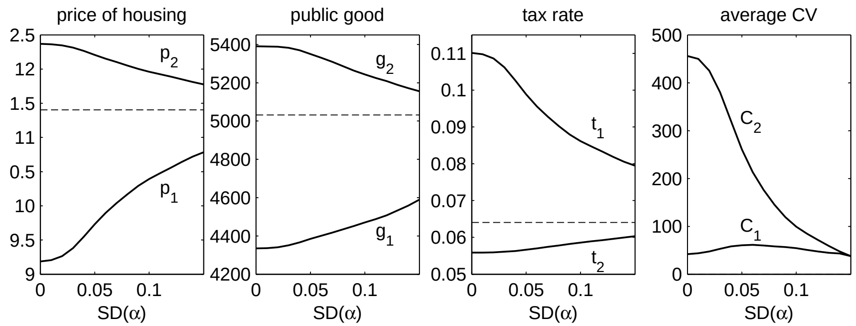

Figure 3.7: Equilibria with changing standard deviation of taste parameter.

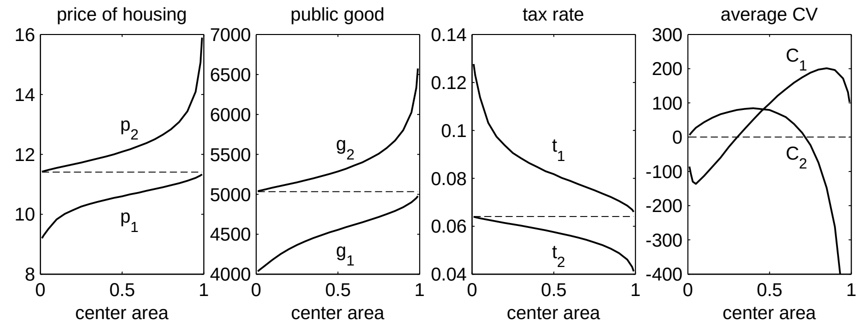

Figure 3.8: Equilibria with changing relative size of the communities.

under the assumption of taste homogeneity are substantially greater than under heterogeneity.

Figure 3.7 shows the equilibrium values for different levels of taste heterogeneity measured by the standard deviation $SD(\alpha)$, leaving the mean of tastes constant. The horizontal axes cover a range from $SD(\alpha) = 0$ (taste homogeneity) to a maximum $SD(\alpha) = 0.143$[22] including the calibrated case

from the one in the case of taste heterogeneity as the pivotal voter varies with the taste parameter. However, this difference is numerically negligible if the constant taste parameter equals the mean of the heterogeneous tastes.

[22]Given the mean $E\alpha = 0.17$, $SD(\alpha) = 0.143$ is the maximal standard deviation that preserves the bell-shaped form of the beta distribution. Higher values lead to a u-shaped

$SD(\alpha) = 0.11$. The picture reveals that the equilibrium converges towards the values of the symmetric case, indicated by the dotted lines. This is explained by the fact that with increasing taste variance, the population is more and more segregated by taste rather than income. This result suggests that taste heterogeneity is able to lower the negative distributional effects of decentralized tax authority. Figure 3.7 (far right) shows the corresponding change of the average compensating variation. While the average CV in the poor center community is almost unaffected by the amount of taste heterogeneity in the population, it falls sharply in the rich periphery. The fraction of households in the periphery which would prefer harmonized taxes falls accordingly from 100% ($SD(\alpha) = 0$) to 65% ($SD(\alpha) = 0.143$).

Up to now, the physical land size was given by the calibration. However, the relative size of the two jurisdictions is likely to influence the equilibrium values and is therefore investigated in detail. The influence of the relative land size is reported in Figure 3.8. The housing price and the public good provision in both communities increase with the physical size of the center community (community 1). Recall that the high-tax community is called the center by convention. The low-tax community shows lower public good provision and lower housing prices than the low-tax community throughout all possible partitions of land between the two communities. Furthermore, the average income in the low-price community is always lower than in the high-price community. The order of community characteristics is hence not affected by the relative land size. Not surprisingly, the equilibrium values of a community that virtually covers the whole area ($L_1 = 1$ or $L_1 = 0$, hence $L_2 = 1$) equal the values of one unified community, marked by the dotted lines. The equilibrium values in the remaining very small community differ maximally from the values in the case of unified communities. The tax rate in the rich community 2 declines with increasing relative land size in community 1. This shows that the rich community's ability to set low taxes is higher when it is physically small.

The influence of the relative land size on welfare is particularly interesting. Recall that in the calibrated situation ($L_1 = 0.4$), the average household prefers harmonized taxes: average CV is CHF 51 in the center and CHF 84 in the periphery. This result does strongly depend on the relative community size as can be seen in Figure 3.8 (far right). The average CV in the poor center community is negative if this community is small ($L_1 < 0.3$), meaning that the population in the center does on average prefer (higher) local taxes to (lower) harmonized taxes, as they are associated with lower housing prices.[23] Note that it is the poorer part of the population in the poor community that profits most from the local differences. The rich periphery shows a similar picture. Its increased ability to set low taxes when it is small ($L_2 < 0.27$) leads to a negative average CV, i.e. an average preference for local taxation.[24] Note that it is the richer part of the population that profits most from the decentralized tax setting.

distribution.

3.4 Conclusions

This chapter presents a model of an urban area with local income taxes used to finance a local public good. The main assumptions of the model are the following: Households differ in income and taste for housing. The demand for housing and private goods is a linear demand system. The share of housing in the budget of the households is on average declining. The private good can only partly substitute for the publicly provided good.

The existence of a segregated equilibrium is shown in a calibrated two-community model assuming realistic single-peaked distributions for income and taste in housing. The high-tax community shows both lower housing prices and lower public good provision than the low-tax community. The equilibrium features segregation of households by both income and tastes. The emerging segregation pattern is such that rich households prefer the low-tax-high-price community given a subpopulation with equal tastes. As tastes differ across households, this does not lead to a perfect income segregation of the population but to an income distribution in the 'rich' low-tax community that stochastically dominates the income distribution in the 'poor' high-tax community. The model is able to explain the substantial differences of the local income tax level and of average income across communities as observed in Switzerland.

[23] The majority of households in the center prefer decentralized tax setting when $L_1 < 0.26$.

[24] When the rich community is even smaller ($L_2 < 0.11$), the supporters of local taxation is on the majority.

The numerical investigation suggests that taste heterogeneity reduces the distributional effects of local tax differences. The differences of characteristics across communities are maximal when tastes are equal for all households and the population is accordingly perfectly segregated by incomes. These differences decrease with increasing taste heterogeneity as the income segregation of the population becomes more and more diffuse.

The numerical investigation also suggests that the relative size of the individual jurisdictions has great impact on the equilibrium situation. The characteristics of a relatively large community are close to the equilibrium characteristics of a single jurisdiction that covers the whole area. Conversely, the relatively small community differs substantially from the single jurisdiction. The ability of the rich community to set low taxes, for example, is higher when this community is physically small. However, contrary to the findings by Hansen and Kessler (2001a), a tax haven need not be small.

Multi-community models are especially well-suited to study metropolitan areas as it is assumed that the residence choice of a household is made after and independent of the decision of where its members work. This assumption does not seem justifiable on the level of federal states or even countries. Therefore, the results presented in this chapter may only be indicative for the analysis of fiscal federalism on the level of states or countries.

From an empirical perspective, a weakness of the model is its poor explanatory power for the typically observed high housing prices in the highly taxed center of Swiss metropolitan areas. This is due to neglect of the advantages of the central business (and cultural) district for households. The incorporation of distance between residence and business or other activities – hence the incorporation of geography – is an interesting task for future research.

3.A Appendix

The household problem is

$$\begin{aligned}\max_{h,b} U(h,b,g,\alpha) &= \alpha\, ln(h-\beta_h) + (1-\alpha)\, ln(b-\beta_b) + \gamma\, ln(g-\beta_g) \\ s.t. &\quad ph + b \le y(1-t)\,.\end{aligned}$$

This leads to the housing demand

$$h^* = h(t,p,y,\alpha) = \frac{\alpha[y(1-t) - p\beta_h - \beta_b]}{p} + \beta_h\,,$$

the income elasticity of housing

$$\varepsilon = \frac{\partial h^*}{\partial y(1-t)} \frac{y(1-t)}{h^*} = \frac{\alpha y(1-t)}{\alpha[y(1-t) - p\beta_h - \beta_b] + p\beta_h}$$

and the indirect utility function

$$\begin{aligned}V &= \alpha\, ln(\alpha) + (1-\alpha)\, ln(1-\alpha) - \alpha\, ln(p) \\ &+ ln[y(1-t) - p\beta_h - \beta_b] + \gamma\, ln(g-\beta_g)\,.\end{aligned}$$

The marginal rates of substitution in Property 1 are derived by totally differentiating the indirect utility function:

$$\begin{aligned}M_{g,t} &= -\frac{\partial V/\partial t}{\partial V/\partial g} = \frac{y(g-\beta_g)}{\gamma[y(1-t) - p\beta_h - \beta_b]}\,, \\ M_{g,p} &= -\frac{\partial V/\partial p}{\partial V/\partial g} = \frac{h^*(g-\beta_g)}{\gamma[y(1-t) - p\beta_h - \beta_b]}\,, \\ M_{t,p} &= -\frac{\partial V/\partial p}{\partial V/\partial t} = -\frac{h^*}{y}\,.\end{aligned}$$

Differentiation of the MRS w.r.t. income and taste yields Property 2:

$$\frac{\partial M_{g,t}}{\partial y} = -\frac{(g-\beta_g)(p\beta_h + \beta_b)}{\gamma[y(1-t) - p\beta_h - \beta_b]^2}\,, \qquad \frac{\partial M_{g,t}}{\partial \alpha} = 0\,,$$

$$\frac{\partial M_{g,p}}{\partial y} = -\frac{(1-t)(g-\beta_g)\beta_h}{\gamma[y(1-t) - p\beta_h - \beta_b]^2}\,, \qquad \frac{\partial M_{g,p}}{\partial \alpha} = \frac{g-\beta_g}{p\gamma}\,,$$

$$\frac{\partial M_{t,p}}{\partial y} = \frac{(1-\alpha)p\beta_h - \alpha\beta_b}{py^2}\,, \qquad \frac{\partial M_{t,p}}{\partial \alpha} = -\frac{y(1-t) - p\beta_h - \beta_b}{py}\,.$$

The independence of the MRS ratio in Property 3 follows directly:

$$\frac{\partial M_{g,p}}{\partial y} / \frac{\partial M_{g,t}}{\partial y} = \frac{(1-t)\beta_h}{p\beta_h + \beta_b},$$

$$\frac{\partial M_{t,p}}{\partial y} / \frac{\partial M_{t,g}}{\partial y} = \frac{[(1-\alpha)p\beta_h - \alpha\beta_b](g-\beta_g)}{\gamma p(p\beta_h + \beta_b)},$$

where $M_{t,g} = 1/M_{g,t}$.

The locus of indifferent households between community j and i

$$\hat{y}_{ji}(\alpha) = \frac{(p_j\,\beta_h + \beta_b)\, p_i{}^{\alpha}\,(g_j - \beta_g)^{\gamma} - (\beta_b + p_i\,\beta_h)\, p_j{}^{\alpha}\,(g_i - \beta_g)^{\gamma}}{(1-t_j)\, p_i{}^{\alpha}\,(g_j - \beta_g)^{\gamma} - (1-t_i)\, p_j{}^{\alpha}\,(g_i - \beta_g)^{\gamma}}.$$

solves $V(t_j, p_j, g_j, y, \alpha) = V(t_i, p_i, g_i, y, \alpha)$ for y. Alternatively, the locus solves for α:

$$\hat{\alpha}_{ji}(y) = \frac{ln[\frac{y(1-t_j)-p_j\beta_h-\beta_b}{y(1-t_i)-p_i\beta_h-\beta_b}] + \gamma\, ln[\frac{g_j-\beta_g}{g_i-\beta_g}]}{ln(p_j/p_i)}.$$

The locus $\hat{\alpha}_{ji}(y)$ is either strictly increasing and concave in y or strictly decreasing and convex, as can easily be verified by inspecting the first and second derivative

$$\frac{\partial\hat{\alpha}_{ji}}{\partial y} = -\frac{(1-t_j)[p_i\beta_h - \beta_b] - (1-t_i)[p_j\beta_h - \beta_b]}{[y(1-t_j) - p_j\beta_h - \beta_b][y(1-t_i) - p_i\beta_h - \beta_b]\, ln(p_j/p_i)}$$

$$\frac{\partial^2\hat{\alpha}_{ji}}{\partial y^2} = -\frac{\partial\hat{\alpha}_{ji}}{\partial y} \cdot \frac{(1-t_j)[y(1-t_i) - p_i\beta_h - \beta_b] + (1-t_i)[y(1-t_j) - p_j\beta_h - \beta_b]}{[y(1-t_j) - p_j\beta_h - \beta_b][y(1-t_i) - p_i\beta_h - \beta_b]}$$

and provided that all household reach the subsistence level, i.e. $y(1-t) > p\beta_h + \beta_b > 0$, in both communities.

The utility difference between community j and i is

$$V_j(y,\alpha) - V_i(y,\alpha) = -\alpha\, ln(\frac{p_j}{p_i}) + ln[\frac{y(1-t_j) - p_j\beta_h - \beta_b}{y(1-t_i) - p_i\beta_h - \beta_b}] + \gamma\, ln(\frac{g_j - \beta_b}{g_i - \beta_b}).$$

Differentiation of the above expression w.r.t. y and α is used in the proof of Propositions 1 and 2:

$$\frac{\partial(V_j - V_i)}{\partial y} = \frac{1}{y - \frac{p_j\beta_h+\beta_b}{1-t_j}} - \frac{1}{y - \frac{p_i\beta_h+\beta_b}{1-t_i}}, \quad \frac{\partial(V_j - V_i)}{\partial\alpha} = ln(p_i) - ln(p_j).$$

The rate of substitution between tax rate and public good provision a voter faces is derived from totally differentiating the indirect utility function considering the housing market reaction, $dp/dt|_{HD=HS}$ (community subscripts ommited):

$$\left.\frac{dg}{dt}\right|_{dV=0,HD=HS} = \frac{-\frac{\partial V}{\partial t} - \frac{\partial V}{\partial p}\cdot\left.\frac{dp}{dt}\right|_{HD=HS}}{\frac{\partial V}{\partial g}} = M_{gt} + M_{gp}\cdot\left.\frac{dp}{dt}\right|_{HD=HS}$$
$$= \frac{g-\beta_g}{\gamma[y(1-t) - p\beta_h - \beta_b]}\left[y + h^*\left.\frac{dp}{dt}\right|_{HD=HS}\right].$$

The voter's rate of substitution is positive when the price effect on the housing market is not too large:

$$\left.\frac{dg}{dt}\right|_{dV=0,HD=HS} > 0 \quad \text{iff} \quad \left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < \frac{y(1-t)}{ph^*}.$$

The voter's rate of substitution decreases with income

$$\frac{\partial\left.\frac{dg}{dt}\right|_{dV=0,HD=HS}}{\partial y} = \frac{\partial M_{g,t}}{\partial y} + \frac{\partial M_{g,p}}{\partial y}\left.\frac{dp}{dt}\right|_{HD=HS}.$$
$$= -\frac{g-\beta_g}{\gamma[y(1-t) - p\beta_h - \beta_b]^2}\left[p\beta_h + \beta_b + (1-t)\beta_h\left.\frac{dp}{dt}\right|_{HD=HS}\right]$$

if the price effect on the housing market is not too large:

$$\frac{\partial\left.\frac{dg}{dt}\right|_{dV=0,HD=HS}}{\partial y} < 0 \quad \text{iff} \quad \left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < \frac{p\beta_h + \beta_b}{p\beta_h}.$$

Both the condition on the sign of the voter's marginal rate of substitution and the sign of its derivative w.r.t. y are fulfilled if $\left.\frac{dp/p}{d(1-t)/(1-t)}\right|_{HD=HS} < 1$ (see Definition 3) and all households reach the subsistence level.

The compensating variation cv_j is the additional gross income that a household in e.g. community j needs in order to be compensated for a shift from the symmetric (unified) equilibrium, (t_u, p_u, g_u), to the segregated equilibrium, (t_j, p_j, g_j). Solving

$$V(t_u, p_u, g_u; y, \alpha) = V(t_j, p_j, g_j; y + cv, \alpha)$$

for cv yields the compensating variation for a household with income y and taste α in community j:

$$cv_j(y, \alpha) = \frac{[y(1-t_u) - \beta_b - p_u\beta_h](\frac{p_u}{p_j})^{-\alpha}(\frac{g_u-\beta_g}{g_j-\beta_g})^{\gamma} - [y(1-t_j) - \beta_b - p_j\beta_h]}{1-t_j}.$$

The average compensating variation in community j is then computed as

$$CV_j = \frac{1}{n_j}\int_0^1 \int_{\underline{y}_j(\alpha)}^{\overline{y}_j(\alpha)} cv_j(y, \alpha)\, f(y, \alpha)\, dy\, d\alpha\,.$$

Chapter 4

Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland

4.1 Introduction

Fiscal Federalism is often viewed as the natural counterpart of decentralized decision making. Oates (1972) for example argued that local units deciding upon public programs are more likely to trade off costs against benefits if these programs are financed by local taxes. Or as Musgrave (1999, p. 156) pointedly remarks: “To secure an efficient outcome, the provision of public services should be determined and paid for by those who benefit.”

While the virtues of decentralized financial responsibility are uncontested, the resulting tax differentials are highly disputed. Tax differentials can be the consequence of different preferences for the level of locally provided public goods. However, different tax rates can also be the result of different economic resources of the local population, since rich local jurisdictions can raise the same revenue with lower tax rates as poor ones. While the effect of the tax base on tax rates is trivial, the opposite effect is less evident. This chapter addresses the question whether tax differentials across local jurisdictions are not just the consequence but also the cause of differences in local average income.

The theoretical part of this chapter proposes the progressivity of a local income tax as a new theoretical explanation for income segregation of the population. The empirical part studies the community choice of households in Switzerland. Swiss metropolitan areas are a laboratory for federal systems as they are divided into a multitude of communities with extensive political and fiscal autonomy. Switzerland is also unique in that the main local tax is on income rather than on property. The estimated multinomial response models show that rich households are significantly and substantially more likely to move to low-tax communities than poor households.

The theoretical literature on the local provision of local public goods goes back to Tiebout (1956). Tiebout showed that fiscal decentralization leads to an efficient provision of local public goods because people with similar preferences would settle in particular communities and vote for their desired level of public goods provision. Tiebout's result rests on the assumption that households have equal incomes. The location of households and the local provision of public goods when the households differ in incomes was studied by Ellickson (1971), Westhoff (1977) and the literature surveyed in Ross and Yinger (1999).

The *segregation hypothesis* is one of the central propositions in multi-community models in the tradition of Tiebout. Endogenous segregation means that different people choose different locations in equilibrium. While the Tiebout model focuses on preference heterogeneity, Ellickson and Westhoff turned the attention to income as the main dimension of difference. Several mechanisms have been proposed that explain why rich households make different choices from poor households (see Ross and Yinger, 1999, for property tax models and Chapter 2 for income tax models). The nature of the local public good, ranging from a monetary transfer to a non-substitutable pure public good, induces a self-sorting of the population when rich households esteem public goods relatively more than poor households. Another mechanism draws on the income elasticity of housing. If housing expenditures become relatively less important with increasing income, rich households are less concerned about high housing prices than poor households.

The segregation mechanism in this chapter builds on the empirical fact that most income tax schemes are progressive and that local jurisdictions

can often only set the tax level within a given federal tax scheme. This mechanism explains the high priority of tax rates in rich households' decisions through the progressivity the of tax scheme.

The segregation hypothesis of Tiebout type models has been challenged by a series of empirical studies.[1] A first strand of research investigates the equilibrium predictions of multi-community models using data on aggregate community characteristics. Epple and Sieg (1999) and Epple, Romer and Sieg (2001) develop a strategy for estimating the household preference parameters of a full equilibrium model where the local income distribution and local policy variables are simultaneously determined. They show that the differing income quantiles across 92 communities in the Boston area are well explained by the model predictions. Feld and Kirchgässner (2001) regress the share of various income classes in Swiss cantons and main cities on income tax rates. They find a strong negative relationship between the tax rate and the share of rich households. However, their treatment of the generic endogeneity of tax rates by instrumental variables from mainly lagged observations does not solve the problem, as the general equilibrium of tax rates and income segregation is most likely a long-run phenomenon. Rhode and Strumpf (forthcoming) assess the importance of the segregation mechanism in Tiebout type models from a historical perspective. They collected an impressive data set with various measures of heterogeneity in the population over a period of 140 years. Given that the costs of moving dramatically declined during this time, multi-community models predict that the population within local units should have become more homogeneous while the differences across local units have aggravated. They conclude that their data do not support the model predictions on a national scale. For metropolitan areas, however, the observed pattern does not contradict the segregation hypothesis.

The second empirical approach - also used in this chapter - directly targets the location choice of individual households using a multinomial response framework. This approach circumvents the endogeneity problem

[1] The early empirical literature on multi-community models investigated the relationship between local tax differentials, public goods provision and housing prices. Oates (1969) and and a multitude of subsequent studies (surveyed in Ross and Yinger, 1999) strikingly confirm the so-called capitalization hypothesis, which predicts that low taxes and attractive public goods provision should be reflected in high housing prices.

because from the perspective of a single household the community characteristics can be taken as given. Friedman (1981) used a conditional logit model to study the location choice of 682 households among nine residential communities close to San Francisco. Nechyba and Strauss (1998) use the same model to study the choice of over 22'000 households among six school districts in the suburbs of Philadelphia. Both studies show that public expenditures are an important locational factor. The segregation hypothesis needs explicit consideration as household specific variables are not identified in linear conditional logit models (see Section 4.4.1). In need of a variable that depends on both household and community characteristics, Nechyba and Strauss calculate the households' hypothetical consumption of private goods for all communities. This variable depends on after-tax local housing prices under the ad-hoc assumption that households consume the same amount of housing in all communities. They therefore implicitly assume that the price elasticity of housing is zero. This assumption is relaxed by using another ad-hoc specification using community-specific coefficients for household income (see footnote 6). Note that the empirical approach depicted in this paragraph neglects the (long-run) reaction of aggregate community characteristics.

Bayer, McMillan and Rueben (2002) attempt a combination of the two empirical approaches. Following Berry, Levinsohn and Pakes (1995) they first estimate the households' choice of a neighborhood, using community fixed effects and a multitude of interaction effects between household and community characteristics. In a second step they explain the community fixed effects by community characteristics using instrumental variables. The estimation strategies in both steps make use of an explicit general equilibrium model. The predictions of the estimated model therefore adequately take into consideration the (long-run) adjustment of the endogenous aggregate community characteristics.

This study follows the second approach but departs from the previous studies by shifting the focus to assessing the (income) segregation hypothesis. The general locational attractiveness of a community, including local public goods, is considered in community specific fixed effects, thereby avoiding the difficulty of measuring public goods provision. The identification of household effects is drawn on an explicit theoretical multi-community

model. Furthermore, recent econometric developments using simulation methods are applied to consider the spatial structure in the error components.

The chapter is organized as follows: Section 4.2 describes the institutional organization of fiscal federalism in Switzerland. A theoretical model of location choice based on progressive income taxation is proposed in Section 4.3. The econometric model is discussed in Section 4.4, while Section 4.5 describes the data. The empirical results and two policy experiments are presented in Section 4.6. Section 7 draws conclusions.

4.2 Fiscal Federalism in Swiss Metropolitan Areas

Switzerland is an exemplary federal fiscal system. The Swiss federation comprises 26 states, the so-called cantons. The cantons are divided into roughly 3000 communities of varying size and population. All three state levels finance their expenditures essentially by their own taxes and fees. The total tax revenue of all three levels was 93 billion CHF in 2001, of which 46% is imposed by the federation, 32% by the cantons and 22% by the communities.[2] While the federal government is mainly financed by indirect taxes (61% of federal tax revenue) such as the VAT, the cantons and communities largely rely on direct taxes. Income taxes account for 60% of cantonal and 84% of communal tax revenue. In total, 46% of the income tax revenue go to the cantons, 38% to the communities and only 16% to the federal government. Transfers between the three levels are not a major part of the budgets of cantons (23% of total revenue) and communities (14%).

The cantons organize their tax systems autonomously. For example, they decide upon the level of income and corporate taxes and the degree of tax progression. The individual communities in turn can generally set a tax shifter for income and corporate taxes. The communal tax is then the cantonal tax rate multiplied by the communal tax shifter. In some cantons, for example in the Canton of Basel-Stadt before 2001, the individual com-

[2] All figures in this paragraph apply to 2001. Source: Swiss Federal Tax Administration (2002), Öffentliche Finanzen der Schweiz 2001, Neuchâtel: Swiss Federal Statistical Office.

munities also have some freedom in setting the tax scheme. The decisions in the cantons and communities are made by the legislative body and are subject to referendums. Federal and cantonal systems of fiscal equalization limit the tax differences across cantons and across communities within the same canton to some extent, but still leave room for considerable variation.

The above outlined federal system leads to ample differences of income taxes across Swiss communities. For example, for a two-child family with a gross income of 80,000 Swiss francs (CHF) the sum of cantonal and community income tax ranged from 3,6% in the city of Zug to 11,3% in Lauterbrunnen in the Canton of Bern in the year 1997 (see the data sources in the appendix). The federal income tax for this household was 0.7%. With an income of 500,000 CHF a two-child family faced much higher tax rates due to the progressive federal and cantonal tax schemes, namely ranging from 10.9% in Wollerau in the Canton of Schwyz to 28.7% in Onex in the Canton of Geneva. The federal income tax for this household was 9.4%.

The tax differences across communities within a single metropolitan area are smaller but still substantial. Figure 1 shows the community characteristics in the metropolitan area of Basel[3] (data sources in the appendix). In 1990 the Basel area was the third largest Swiss metropolitan area with a total of 406,000 inhabitants. The city of Basel with 178,000 inhabitants, hereafter called the center, is the central business district of the area. The top-left map shows the share of workers commuting to the center. The white area to the north and west of the City of Basel is French and German territory and is not included in this study. The Basel area comprises 38 communities from four cantons: Basel-Stadt, Basel-Land, Solothurn and Aargau. There is great variability in both tax levels and tax schemes. The totalled communal and cantonal income tax rate for a two-child family is depicted in the bottom-left map. The taxes are highest in the center community (9.4 %) and up to 35% lower in the communities around the center.

It is particularly interesting to contrast the local tax rates with the income of the residents. The bottom-right map in figure 1 shows the local share of households with incomes above 75,000 CHF. The map represents to a great extent an inverted picture of the tax rates. The high-tax cen-

[3]Definition of the area according to the Swiss Federal Statistical Office based on Census 1990 data.

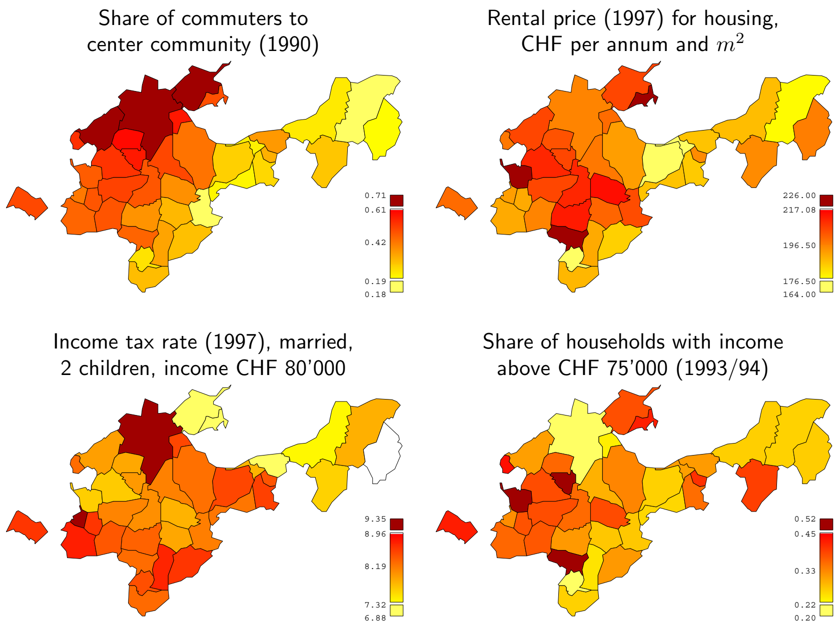

Figure 4.1: Community characteristics in the metropolitan area of Basel.

ter community has the lowest proportion of rich households, whereas the low-tax communities close to the center are populated by many more rich households. The rental prices for housing also seem to be correlated with the tax rates. The low-tax fringe of communities around the center exhibits higher average prices than the center, although communities further away from the center are clearly the lowest-price locations.

4.3 A Model of Location Choice and Local Progressive Income Taxation

The theoretical model describes a metropolitan area with a fixed number J of distinct local jurisdictions, called communities. The political borders of the communities are the outcome of a historical process and thus taken as given. The area is populated by a continuum of heterogeneous households, which differ in incomes. Income is distributed according to a distribution function $f(y) > 0$ with support $[\underline{y}, \overline{y}], \underline{y} > 0, \overline{y} < \infty$. There are three goods in the economy: private consumption b, housing h and a public good g.

The local public expenditures are financed by a tax on the residents' income. The income tax rate $t_j(y)$ in a community j depends on the households' income y. The provision of the public good g is fixed by the state government and, hence, is identical across communities. This assumption is motivated by the observation that the autonomy of local decision making is in fact often strongly limited by state and federal laws.[4] In the case of education expenditures, for example, teachers' salaries as well as class sizes are regulated by the cantons.

The price for housing p_j in community j is determined on a competitive housing market. Hence, The communities are fully characterised by their local income tax level and their local price for housing. A household can move costlessly and chooses the community maximizing its utility as place of residence.

This chapter focuses on the households' location decision and does not develop a full general equilibrium model. A complete model includes the description of the housing supply function, the production function of the public good and the equilibrium concept. Chapter 1 (Kapitel 1) presents simulated equilibria in a fully specified complete model

[4]The exogenous determination of public goods provision substantially simplifies the model and turns the focus to income segregation induced by local taxation. A comprehensive model of local provision of local public goods would have to endogenize the provision of local public goods. However, this greatly complicates the analysis without providing qualitatively different results for location choice and income segregation. In addition, the more general approach makes it intractable to consider progressive tax schemes, which are crucial in the empirical investigation. See Chapter 2 for a discussion of the technical problems and Chapter 3 for the properties of a model with endogenous local public goods provision financed by local income taxes.

4.3.1 Household Preferences

The preferences of the households are described by a utility function [5]

$$U(h, b) \ ,$$

where h is the consumption of housing and b the consumption of the private good. The utility function is assumed to be strictly increasing, strictly quasi-concave and twice continuously differentiable in h and b.

Households face a budget constraint:

$$ph + b \leq y_d = y[1 - t \cdot r(y)] \ ,$$

where p is the price of housing. The price of the private good is set to unity. The disposable income y_d depends on the local income tax shifter $t > 0$ and the exogenous tax rate structure $r(y)$. The tax scheme $r(y) > 0$ is assumed to be continuous in y. The (average) tax rate $t_j(y) = t \cdot r(y)$ is smaller than the marginal tax rate $\partial(t\, r(y)\, y)/\partial y = t[r + y\, r'(y)]$ and both are assumed to lie in $[0, 1)$.

Maximisation of the utility function with respect to h and b subject to the budget constraint yields the housing demand function $h^* = h(p, y_d) = h(t, p\,; y)$, the demand for the private good $b^* = y(1 - t\, r) - ph(t, p\,; y)$, and the indirect utility function

$$V(t, p\,; y) := U(h^*, b^*) \, . \tag{4.1}$$

Property 1 is a trivial result of the strictly increasing nature of the utility function and is derived by applying the implicit function theorem and the envelope theorem:

Property 1 (MRS between community characteristics)

$$M(t, p, y) := \left.\frac{dt}{dp}\right|_{dV=0} = -\frac{\partial V/\partial p}{\partial V/\partial t} = -\frac{h^*}{y \cdot r(y)} < 0 \, .$$

[5]The public good does not explicitly enter the utility function because it does not affect the following considerations as it is assumed to be constant across communities. This assumption is relaxed in Section 4.3.4.

The marginal rate of substitution (MRS) between community characteristics reflects a household's trade-off between taxes and housing prices. Property 1 simply follows from the fact that households dislike both high taxes and high housing prices. A household can therefore be compensated for a tax increase by a decline in housing prices and vice-versa.

The following two assumptions about the form of the indirect utility function generate the segregation by income.

Assumption 1 (Income elasticity of housing)

$$\varepsilon_{h,y_d} := \frac{\partial h^*}{\partial y_d} \frac{y_d}{h^*} \leq 1 \quad \textit{for all } y_d \textit{ and } p.$$

Assumption 1 means that housing is a normal good, i.e. the elasticity of housing with respect to disposable income is smaller or equal to unity. This implies that the share of housing in the household's budget decreases with after-tax income.

Assumption 2 (Progressive taxation)

$$\frac{\partial r(y)}{\partial y} \geq 0 \quad \textit{for all } y.$$

Assumption 1 states that the income tax scheme is proportional or progressive.

Property 2 (Relative preferences)
If Assumptions 1 and 2 hold and if and only if one of them holds with strict inequality, then

$$\frac{\partial M}{\partial y} = \left[1 - \frac{\partial h^*}{\partial y_d} \frac{y_d}{h^*} \frac{\partial y_d}{\partial y} \frac{y}{y_d}\right] \frac{h^*}{y^2 r(y)} + \frac{\partial r(y)}{\partial y} \frac{h^*}{y r^2(y)} > 0$$

for all y, t and p.

Proof: Assumption 1 states that $(\partial h^*/\partial y_d)(y_d/h^*) \leq 1$. The assumptions about the relation and the bounds of the average and the marginal tax rate guarantee that $(\partial y_d/\partial y)(y/y_d) = [1 - t\, r(y) - t\, y\, r'(y)]/[1 - t\, r(y)]$ lies in $[0, 1]$. If Assumption 2, $\partial r(y)/\partial y > 0$ is strictly satisfied, both addends of $\partial M/\partial y$ are strictly positive. If Assumption 2 is not strictly satisfied, $\partial r(y)/\partial y = 0$, and Assumption 1 is strictly satisfied, $(\partial h^*/\partial y_d)(y_d/h^*) < 1$, then the second addend is zero and the first addend is strictly positive. If $\partial r(y)/\partial y = 0$ and $(\partial h^*/\partial y_d)(y_d/h^*) = 1$ then both addends are zero. □

Property 2 states that the MRS between local tax levels and housing prices increases monotonically with income. This means that rich households have a relatively stronger preference for low taxes than poor households. Property 2 explains why rich households make different location decisions than poor households. It is therefore the central condition giving rise to income segregation. Westhoff (1977) called the analogous assumption 'relative preference assumption'. It is also called the single-crossing condition. In this model, relative preferences are either caused by the progressive tax scheme, the income elasticity of housing below unity or a combination of both. As will become apparent in Section 4.4, Property 2 plays a key role in the identification of tax rate effects in random utility maximization models of location choice.

4.3.2 Location Choice

A household with income y chooses the community which maximizes its utility. Hence, given the set of community characteristics (t_j, p_j) for $j \in C = (1, ..., J)$, a household prefers community j if and only if

$$V(t_j, p_j\,; y) \geq V(t_i, p_i\,; y) \quad \text{for all } i\ . \tag{4.2}$$

The following propositions describe the allocation of households to communities when all communities are populated and exhibit different characteristics.

Proposition 1 (Order of community characteristics)
If all communities are populated and exhibit different community characteristics, then communities with higher housing prices impose lower income tax rates.

Proof: Suppose the opposite, i.e. that one community exhibits both lower prices and lower taxes. Then all households would prefer that community for the same reason that lead to Property 1. This is a contradiction. □

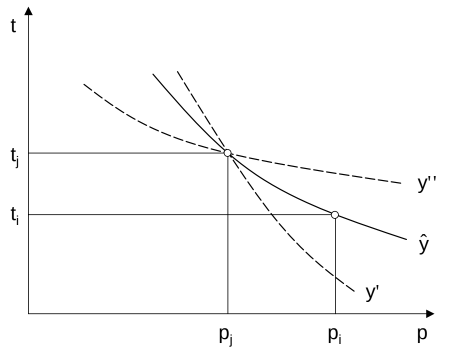

Figure 4.2: Indifference curves in the (t, p) space.

Proposition 2 (Perfect income segregation)
If the relative preference property holds and all communities are populated and exhibit different community characteristics, then all households choosing a community with lower taxes are richer than all households choosing a community with higher taxes.

Proof: The proof proceeds in three steps. Firstly, it is shown that there is a 'border' household in a comparison of two communities. Secondly, income segregation is shown in a two community case. Thirdly, the result is extended to more than two communities.

(1) Define $V_j(y) := V(t_j, p_j, y)$ as a household's utility in j and $V_i(y) := V(t_i, p_i, y)$ in i. Let the household with income y' prefer j to i, hence $V_j(y') - V_i(y') \geq 0$ and a household with income y'' prefer i: $V_j(y'') - V_i(y'') \leq 0$. From the continuity of V in y follows the continuity of $V_j(y) - V_i(y)$ in y. The intermediate value theorem states that there is at least one $\hat{y}$ between y' and y'' s.t. $V_j(\hat{y}) - V_i(\hat{y}) = 0$. This household is called the border household.

(2) This part uses Figure 4.2. The figure shows the indifference curves in the (t, p)-space for three different income levels $y' < \hat{y} < y''$. The indifference curves represent all (t, p) pairs that households consider to be as good as community j's (t_j, p_j)-pair. Households prefer pairs south-west of the indifference curve to (t_j, p_j). Note that the indifference curves are decreasing in the (t, p)-space (Property 1). Note also that, due to Property 2, they become flatter as income rises. Imagine a community i, characterized by (t_i, p_i), $p_i > p_j$ and $t_i < t_j$, where household $\hat{y}$ is indifferent to j. All richer households, e.g. y'', prefer the low-tax community i to j and all poorer households, e.g. y', prefer the low-price community j.

(3) The proposition implies that $[\underline{y}, \overline{y}]$ is partitioned into J non-empty and non-overlapping intervals $I_j = \{y | \text{household with income } y \text{ chooses } j\}$. Suppose the opposite, i.e. y' as well as y'' prefer community j, but an y''', $y' < y''' < y''$ strictly prefers community i. It follows from step 1 that there is an $\hat{y}$, $y' \leq \hat{y} < y'''$. Step 2 implies that $y'' > \hat{y}$ strictly prefers i to j, which is a contradiction. □

Proposition 2 claims that any community is populated by a single and distinct income class or more fundamentally that rich households choose different communities than poor households. This proposition is assessed in the empirical part of this chapter.

Proposition 3 (Non-existence of income segregation)
If the local income tax rate is proportional and the household preferences are homothetic, then rich households choose the same communities as poor households.

Proof: Neither Assumption 1 nor 2 are satisfied with strict inequality. Therefore, Property 2 does not hold and the indifference curves in Figure 4.2 coincide. Hence, all households are, independently of their income, either indifferent between all communities or all prefer the same community. □

Proposition 3 shows that Property 2 is a necessary condition for income segregation. There is no systematically different behavior of rich and poor households in the absence of a 'screening device' such as progressive taxation and/or nonproportional housing demand.

4.3.3 Adding Taste Heterogeneity

So far, it has been assumed that households with identical preferences differ by income. This section extends the basic model by letting the households differ in both income $y \in [\underline{y}, \overline{y}]$, $0 < \underline{y}, \overline{y} < \infty$, and a parameter $\alpha \in [0, 1]$ describing their taste for housing. Income and taste are jointly distributed according to the density function $f(y, \alpha) > 0$.

The housing preference enters the utility function $U(h, b; \alpha)$ and the indirect utility

$$V_j = V(t_j, p_j\,; y, \alpha) = U(h_j^*, b_j^*; \alpha)\,. \tag{4.3}$$

Households with a larger preference parameter α are assumed to spend, ceteris paribus, more on housing than households with a small α. The housing demand function thus increases with α:

Assumption 3 (Housing taste)

$$\frac{\partial h^*}{\partial \alpha} = \frac{\partial h(t, p; y, \alpha)}{\partial \alpha} > 0 \quad \textit{for all } t,\ p,\ y \textit{ and } \alpha.$$

This specification of preference heterogeneity preserves income segregation within a subpopulation with identical preferences. Moreover, segregation of preferences emerges:

Proposition 4 (Preference segregation)
Consider a subpopulation with equal income y. If all communities are populated and exhibit different community characteristics, then all households choosing a community with higher housing prices have a weaker taste for housing than all households choosing a community with lower housing prices.

Proof: The proof is analogous to Proposition 2 using the counterpart to Property 2,

$$\frac{\partial M}{\partial \alpha} = -\frac{\partial h^*}{\partial \alpha}\frac{1}{y \cdot r(y)} < 0\,. \quad \square$$

Simultaneous heterogeneity by incomes and tastes leads to a more realistic pattern of household segregation in a metropolitan area. Although income groups tend to gather, the segregation is not perfect. Figure 4.3 shows the resulting allocation of household types to communities. The households on the borders are indifferent between neighboring communities j. Community 1 with the lowest housing prices is populated by the poorest households with strong taste for housing, while the richest households with low housing taste are situated in community J with the lowest tax rate and the highest housing price. However, rich households with strong taste for housing prefer lower-priced communities and poor households with weak taste for housing group with relatively rich households in the lower-tax communities.

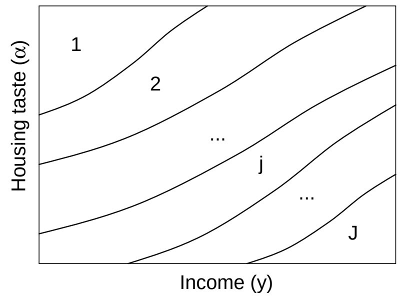

Figure 4.3: Simultaneous income and preference segregation. The areas denoted by $j = 1, ..., J$ show the attributes of the households that prefer community j.

4.3.4 Adding Intrinsic Community Attractiveness

In reality, communities differ in much more than tax levels and housing prices. Factors such as vicinity of the central business district, cultural activities, shopping opportunities, climate and landscape can have an influence on a household's choice of residential location. In addition, the provision of local public goods such as schools, though exogenously set here, can differ across communities.

A straightforward way to incorporate these factors is to add a community specific constant k_j to the (indirect) utility function:

$$V_j = V(t_j, p_j, k_j\,; y, \alpha) = U(h_j^*, b_j^*) + k_j\,. \tag{4.4}$$

This extension does not fundamentally change the logic of the model. Propositions 2, 3 and 4 still hold because the additive separable specification preserves properties 1 and 2. Proposition 1, however, does not hold any more as one can imagine communities so attractive that they can attract households even when they show both higher taxes and higher housing prices. This is usually observed for the center communities in Swiss metropolitan areas.

4.3.5 A Benchmark Case

This section presents the model with a specified utility function for homothetic preferences. Income segregation is therefore solely induced by the progressivity of the tax scheme. The derived indirect utility function will serve as a benchmark in the empirical study.

Household preferences are described by a Cobb-Douglas utility function

$$U(h, b, g; \alpha) = \alpha\, log(h) + (1 - \alpha)\, log(b) + \gamma\, log(g)\,.$$

The resulting demand for housing

$$h^* = h(t, p\,; y, \alpha) = \alpha y[1 - t\, r(y)]p^{-1}$$

increases with α. The parameter $\alpha \in (0, 1)$ can therefore be seen as a measure for housing taste as defined in Section 4.3.3.

The indirect utility function in community j is

$$V_j = V(t_j, p_j\,; y, \alpha) = k - \alpha\, log(p_j) + log(y) + log[1 - t_j\, r(y)]\,, \qquad (4.5)$$

where $k = \alpha\, log(\alpha)+(1-\alpha)\, log(1-\alpha)+\gamma\, log(g)$. The intrinsic attractiveness of the place as outlined in Section 4.3.4 is considered by using a community specific constant k_j.

4.4 The Econometric Model

The empirical part of this study aims to establish factors that determine a household's choice of the residential community in a metropolitan area. The location choice in the theoretical model of the previous section leads naturally to a multinomial response model based on random utility maximization (RUM). Multinomial response models are closely connected with McFadden's (1974, 2001) seminal work on 'economic choices'. The subsequent presentation draws upon Train (2003).

The choice of one of many unordered alternatives is driven by a latent variable, often interpreted as indirect utility. The indirect utility V_{nj} of a household n in a community j is the sum of a systematic and a stochastic part

$$V^*_{nj} = V_{nj} + \varepsilon_{nj}\,,$$

where n indicates the household and j the community. V_{nj} is a deterministic function of observable household and community characteristics and ε_{nj} is a household and community specific error term.

A household n chooses community j among the choice set $C = (1, ..., J)$ as its place of residence if it offers the highest value of indirect utility, i.e.

$$V^*_{nj} \geq V^*_{ni} \quad \text{for all } i \in C\,. \qquad (4.6)$$

This is equivalent to equation (4.2) in the theoretical model.

4.4.1 Functional Form and Identification

The indirect utility function (equation 4.4) in the theoretical part guides the choice of systematic factors in the indirect utility function

$$V_{nj} = V(t_{nj}, p_j, k_j, y_n, a_n)\,, \qquad (4.7)$$

where t_{nj} is the income tax rate of household n in community j, p_j is the housing price in community j, k_j indicates further community specific dimensions of attractiveness, y_n is household income and a_n indicate further household characteristics.

From the point of view of an individual household, the community characteristics are exogenous, although they are the (long-term) aggregate of the agents' decisions. Therefore the household's location decision is optimal given the community characteristics. In the theoretical model this is established by the assumption that there is a continuum of households, i.e. that a single household is 'small' and does not influence the equilibrium.

For the empirical implementation the functional form of the deterministic part (equation 4.7) of the latent variable needs to be specified. Starting point is the indirect utility function (4.5) from the benchmark case presented in Section 4.3.5,

$$V_{nj} = \beta_0\, k_j + \beta_1\, log(1 - t_{nj}) + \beta_2\, log(p_j) + \beta_3\, log(y_n)\,,$$

where k_j and β_0 can be vectors. Note that the parameters are identical across the alternatives.

The theoretical model offers two mechanisms which explain why rich households move to different communities than poor households: progressive taxation and income elasticity of housing below unity. Two interaction terms are added to allow for the latter segregating mechanism:

$$V_{nj} = \beta_0\, k_j + \beta_1\, log(1 - t_{nj}) + \beta_2\, log(p_j) + \beta_3\, log(y_n) \\ + \beta_4\, log(1 - t_{nj})\, log(y_n) + \beta_5\, log(p_j)\, log(y_n)\,. \tag{4.8}$$

The implied MRS between tax rate and housing price satisfies Property 2 (relative preferences) even in the case of proportional taxes

$$\left.\frac{\partial M}{\partial y}\right|_{t_{nj}=t_j} = \frac{(1-t_j)(\beta_1\,\beta_5 - \beta_2\,\beta_4)}{p_j[\beta_1 + \beta_4\, \log(y_n)]^2\, y_n} > 0 \tag{4.9}$$

if $\beta_1\,\beta_5 > \beta_2\,\beta_4$. As one expects $\beta_1 > 0$ and $\beta_2 < 0$ for any household this is guaranteed by $\beta_4 > 0$ and $\beta_5 > 0$. This means that the effect of the tax rate increases with income while the effect of the housing price decreases. As is shown in the proof of Property 2, the progressive tax scheme reinforces this segregation mechanism.

The idea of heterogenous tastes for housing in Section 4.3.3 is applied by substituting the constant housing price effect β_2 with a household dependent effect $\beta_6 + \beta_7\, a_n$:

$$V_{nj} = \beta_0\, k_j + \beta_1\, log(1 - t_{nj}) + \beta_6\, log(p_j) + \beta_3\, log(y_n) \\ + \beta_4\, log(1 - t_{nj})\, log(y_n) + \beta_5\, log(p_j)\, log(y_n) + \beta_7\, log(p_j)\, a_n\,,$$

where a_n and β_7 can be vectors.

The *level* of the indirect utility function is not identified as the agents only care about the differences between alternatives. Consequently, factors that shift the indirect utility of all alternatives in the same way are not identified; hence β_3 cannot be estimated. This leads to the identified indirect utility function:

$$V_{nj} = \beta_0\, k_j + \beta_1\, log(1 - t_{nj}) + \beta_6\, log(p_j) \\ + \beta_4\, log(1 - t_{nj}) log(y_n) + \beta_5\, log(p_j)\, log(y_n) + \beta_7\, log(p_j)\, a_n\,. \tag{4.10}$$

Note that the *scale* of the indirect utility function will be arbitrarily set by the specification of the error term.

The community characteristics k_j may be imprecisely measured or not observable. It is therefore advantageous to include community fixed effects which capture all unobserved dimensions of intrinsic community attractiveness. However, the effect of an observed community characteristic cannot be distinguished from the fixed effect of this community and is thus not identified. The identified fixed effects specification is:

$$V_{nj} = \delta_j + \beta_1\, log(1 - t_{nj}) + \beta_4\, log(1 - t_{nj}) log(y_n) \\ + \beta_5\, log(p_j)\, log(y_n) + \beta_7\, log(p_j)\, a_n\,, \tag{4.11}$$

where the community-specific constant δ_j is identified by setting the constant of an arbitrary community to zero. Note that the effect of the tax rate t_{nj} can still be estimated because it depends on both the community j and the household n.

4.4.2 Modelling the Stochastic Part

So far, the stochastic element ε_{nj} of household n's utility in community i has not been discussed. The stochastic part stands for all factors of community choice that are hidden from the researcher but known to the household. It therefore represents all *unobserved* factors such as more detailed socio-demographic information about the household as well as all *unobservable* factors such as the household members' attachment to a certain place. There is very little theoretical guidance that would help to model the stochastic term. Several specifications are therefore used and compared in the empirical analysis.

The first specification assumes that the error terms follow independently and identically an extreme value distribution. The cumulative distribution function is

$$F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}}\,.$$

This leads to the *conditional logit* model.[6] The probability that household n chooses community j is

$$P_{nj}(\theta) = \frac{e^{V_{nj}}}{\sum_{i=1}^{J} e^{V_{ni}}}\,, \tag{4.12}$$

[6]The conditional logit model is also called *multinomial logit* model. Modern treatises on multinomial response use the notion of multinomial logit for a specification in which the slope parameter β_j depends on the alternative j. These alternative specific parameters are difficult if not impossible to derive from economic choice behavior.

where V_{ni} is the deterministic part of the utility of household n in community i and $\theta = \beta$ is the set of parameters to be estimated. The independence of the error term across the alternatives is a strong assumption. It implies that a household's stochastic, i.e. unobserved, preference for a certain community is fully independent of its stochastic preference for other communities. The strong and unpleasant consequences of this assumption are discussed in the literature as *independence of irrelevant alternatives* (IIA).

The *nested logit* model is a generalization of the conditional logit model that avoids IIA by allowing a specific pattern of correlations across the error terms (see McFadden, 1984). The vector of all community specific error terms $\varepsilon_n = (\varepsilon_{n1}, ..., \varepsilon_{nJ})$ follows the generalized extreme value distribution (GEV) introduced by McFadden (1978):

$$F(\varepsilon_n) = e^{[-\sum_{k=1}^{K} \left(\sum_{i \in C_k} e^{-\varepsilon_{ni}/\lambda_k}\right)^{\lambda_k}]}.$$

The choice set $C = (1, ..., J)$ is divided into K mutually exclusive subsets C_k, called nests. The unobserved portions of utility ε_{ni} are correlated within the same nest k and independent across nests. The parameter λ_k captures the correlation within nest k. $1 - \lambda_k$ can be used as an indication of correlation, but the link is more complicated (see McFadden, 1978). The extreme case $\lambda_k = 1$ means that there is no correlation within nest k. The nested logit model is consistent with random utility maximization if (but not only if; see Börsch-Supan, 1990) $\lambda_k \in [0, 1]$. Setting all λ_k to unity leads to the conditional logit model. The probability that household n chooses community j is

$$P_{nj}(\theta) = \frac{e^{V_{nj}/\lambda_l} \left(\sum_{i \in C_l} e^{V_{ni}/\lambda_l}\right)^{\lambda_l - 1}}{\sum_{k=1}^{K} \left(\sum_{i \in C_k} e^{V_{ni}/\lambda_k}\right)^{\lambda_k}}, \qquad (4.13)$$

where l is the nest of community j and $\theta = (\beta, \lambda)$.[7] The nested structure of the error term can be looked at as the result of a two-stop choice: households choose a certain nest first and afterwards an alternative within the nest. In the empirical study the first step is naturally the decision whether to stay in the center community or to move to a community in the periphery. Households with a large unobserved preference for a community in the periphery therefore also have a higher preference for all other communities in the periphery. In this case the center community is a nest on its own, called a degenerate nest with $\lambda_k = 1$. This nested structure can be considered as a simple form of a spatial correlation pattern.

[7]Note that this form of the likelihood function is directly derived from the random utility model and the generalized extreme value distribution. Some software packages, e.g. the `nlogit` command in Stata, and textbooks, e.g. Greene (2003), use a slightly different likelihood function in their implementation of nested logit. These likelihood functions are not consistent with random utility maximisation. See Hensher and Greene (2002) for a critical discussion. Stata offers a revised command `nlogitrum` (see Heiss, 2002) which correctly implements a nested logit model. This command is, however, not able to deal with degenerate nests and a full set of alternative fixed effects.

The *multinomial probit* model enables a more flexible specification of the error term compared to the previous two models. The vector of error terms across alternatives is assumed to follow a J-variate normal distribution

$$\varepsilon_n \sim N(0, \Omega),$$

where Ω is the $J \times J$ variance-covariance matrix. This general form allows for all possible correlation patterns across the unobserved part of utility. This flexibility, however, comes at a price: the estimation of multinomial probit models is numerically demanding (see Section 4.4.3) and the general variance-covariance needs to be restricted for both theoretical and practical reasons. Due to the fact that the agents only care about the utility differences across alternatives, Ω needs normalizing and only a maximum of $[(J-1)J/2] - 1$ parameters can be estimated compared to the $J(J+1)/2$ distinct elements in Ω (see Train 2003). In the case of e.g. 17 alternatives there are still 135 parameters to be estimated. These parameters are in practice hardly identified. This study uses a very parsimonious specification of Ω. Following Bolduc (1992) and Bolduc, Fortin and Gordon (1997),[8] the alternative specific error terms follow a first order spatial autoregressive process (SAR)[9]

$$\varepsilon_n = \rho W \varepsilon_n + \xi_n,$$

where $\xi_n \sim N(0, I)$ and $\rho \in (-1, 1)$ is a parameter to be estimated. W is an exogenous $J \times J$ weighting matrix where the weight w_{ji} is a decreasing

[8]Bolduc, Fortin and Fournier (1996) present one of the rare applications of SAR in multinomial response models. They use a slightly different specification and mix the multivariate normal SAR process with an extreme value distribution.

[9]See Anselin and Florax (1995) for a general treatise of SAR processes.

function of the distance d_{ij} between community j and i

$$w_{ji} = \frac{1/d_{ji}}{\sum_{s=1}^{J} 1/d_{js}}$$

and satisfies $w_{ji} = w_{ij}$, $w_{ii} = 0$ and $\Sigma_s w_{is} = 1$ by construction. The variance-covariance of the error term can be derived as

$$\Omega(\rho) = (I - \rho W)^{-1}(I - \rho W)^{-1}$$

because $\rho \in (-1, 1)$ guarantees the nonsingularity of $(I - \rho W)$ (see Berman and Plemmons, 1994, p.133). The probability that household n chooses community j is

$$P_{nj}(\theta) = Prob\left[\varepsilon_1 - \varepsilon_j > V_{nj} - V_{n1}, \ldots, \varepsilon_J - \varepsilon_j > V_{nj} - V_{nJ}\right], \qquad (4.14)$$

where $\theta = (\beta, \rho)$. The above spatial pattern means that households with a strong unobserved taste for a certain community also like other communities geographically close to that community.

4.4.3 Estimation

The conditional and nested logit models are estimated using maximum likelihood (ML) and full information maximum likelihood (FIML) respectively. The log likelihood function is

$$log\,\mathcal{L}(\theta) = \sum_{n=1}^{N}\sum_{j=1}^{J} z_{nj}\, log\, P_{nj}(\theta)\,,$$

where $z_{nj} = 1$ if the household n chooses community j and $z_{nj} = 0$ otherwise. The choice probabilities P_{nj} of the conditional logit and nested logit model are defined in equations (4.12) and (4.13), respectively. The maximum likelihood estimator $\hat{\theta} = (\hat{\beta}, \hat{\lambda})$ is consistent, asymptotically efficient and normally distributed.

The multinomial probit model is estimated with maximum simulated likelihood (MSL, see Hajivassiliou and Ruud, 1994). The calculation of the likelihood requires the integration of a 16-variate normal distribution. As there is no analytic solution to this problem numerical integration routines

or simulation methods are applied. A standard method is the Geweke-Hajivassiliou-Keane GHK choice probability simulator (see Geweke, Keane and Runkle, 1994 and Börsch-Supan, and Hajivassiliou, 1993). GHK simulates the choice probabilities P_{nj} in equation (4.14) by recursively drawing from univariate normal distributions. The number of draws R determines the quality of the approximation. This study uses $R = 1000$ pseudo-random draws in each dimension. The properties of the MSL estimator $\hat{\theta} = (\hat{\beta}, \hat{\rho})$ are equivalent to standard ML if the number of draws R grows faster than $\sqrt{N}$ (see e.g. Train, 2003).

All estimations are performed with the author's own programs in MATLAB.[10] The Newton-Raphson algorithm with the Broyden-Fletcher-Goldfarb-Shanno method (BFGS) for updating the hessian matrix was used for numerical maximization. All parameters, including the coefficients of the correlation structure have been appropriately scaled during optimization. The numerically demanding estimation of the multinomial probit model runs around 70 hours on a Sun Fire V880.

4.5 Data

The empirical investigation is based on non-public household data from the tax administration of the Canton of Basel-Stadt. The data contain information of all households in the city of Basel that moved within the city or from the city to a community in the periphery in the year 1997 .

The decision maker in the theoretical model is a household. Households are operationalized as all persons that moved from a common old address to a common new address: families in a narrower sense, married and unmarried couples as well as people who simply share a flat.[11]

The choice set of these households consists of roughly 3000 Swiss communities and the communities in neighboring France and Germany and in principal the whole rest of the world. However, from both a theoretical and a practical point of view this potential choice set is not the relevant one

[10] A MATLAB toolbox with programs for conditional logit, nested logit, multinomial probit and mixed logit models is available from the author on request.

[11] Married couples that move from single households into a common flat are also treated as one household. Unmarried couples that start living together at the new address are treated as independent households.

in the analysis. Tiebout type models of location choice are only suitable for narrow metropolitan areas.[12] Moreover, the econometric methods used are numerically unfeasible for large choice sets. The analysis is therefore restricted to the city of Basel and a circle of the 16 most integrated communities around it.[13] This leaves 7,872 households with 11,540 members in the data set. The communities belong to three different cantons, Basel-Stadt (BS), Basel-Land(BL) and Solothurn (SO) and thus exhibit great variability in tax levels and tax schemes.

Some information on the communities in the choice set is also used in the analysis. As this study uses community fixed effects, little effort was devoted to finding variables describing community attractiveness. The following enumeration describes the variables used. See the appendix for a detailed description of the variables and the data sources.

- *Income* (household specific): Total gross income of all household members according to the last tax assessment before moving.[14]
- *Marital status* (household specific): Marital status of the primary earner.
- *Children* (household specific): Number of under-age children.
- *Tax rate* (household and community specific): Tax rate for totalled cantonal (state) and communal income taxes. It reflects community/state specific tax deductions, community/state specific progressive tax schemes and community specific tax shifters and thus depends on household income as well as on marital status and children. The hypothetical tax rate is computed for any household as well as any of the 17 communities in the choice set.
- *Rent* (community specific): Average offer price per m^2 for a rented flat.
- *Distance* (community specific): Distance in km^2 between a community and the central business district.

[12]Tiebout type models ignore the location of the work place. When households decide upon their place of residence on a national or global scale, job opportunities are naturally very important. In narrow metropolitan areas, however, it is reasonable to assume that any community is a feasible place of residence for households whose members are working in the central business district.

[13]These communities are defined as all communities where more than 36% of the working population is commuting to the center community (Census 1990). This admittedly arbitrary cutting off point leads to a well-shaped geographic area and a tractable number of choice opportunities. The five smallest communities are omitted as they are not covered in the tax scheme data. Changing the choice set did not qualitatively change the results of the analysis.

[14]The relevant gross income would be the gross income after moving, which is not available. Income before moving is a good proxy if a household's decision to move does not coincide with a change in its income.

Table 4.1: Characteristics of movers from the center community in 1997.

	households moved in	mean income	median income	children	distance	rent	tax mid income†	tax high income†
Whole Area	7872	61,612	54,449	0.32		206	8.05	22.52
City of Basel (BS)	6370	59,334	52,328	0.32	0	197	9.36	26.41
Periphery	1502	71,271	61,874	0.31		207	7.97	22.30
- Binningen (BL)	165	73,405	60,106	0.19	2.5	205	7.88	21.80
- Birsfelden (BL)	98	52,351	52,033	0.23	3.2	200	8.51	23.19
- Bottmingen (BL)	43	76,376	74,131	0.37	4.1	206	7.98	22.08
- Allschwil (BL)	251	69,302	63,138	0.30	4.6	207	7.94	21.77
- Münchenstein (BL)	92	58,962	54,567	0.29	4.9	198	8.13	22.26
- Oberwil (BL)	80	77,048	64,702	0.21	5.4	211	7.66	21.05
- Riehen (BS)	280	83,950	72,428	0.39	5.6	206	6.88	21.77
- Muttenz (BL)	114	63,333	56,688	0.35	5.7	192	8.24	22.66
- Bettingen (BS)	9	69,978	67,177	0.11	6.2	220	7.20	20.86
- Reinach (BL)	151	72,242	61,992	0.25	6.5	212	8.04	22.53
- Arlesheim (BL)	56	57,601	56,688	0.21	7.2	215	7.81	21.88
- Therwil (BL)	46	91,735	79,672	0.54	7.3	207	8.11	22.73
- Biel-Benken (BL)	18	88,610	72,350	0.28	7.8	226	7.64	20.87
- Aesch (BL)	57	62,968	53,506	0.35	9.5	213	8.33	23.33
- Ettingen (BL)	24	61,541	65,999	0.38	10.1	197	8.40	23.54
- Hofst.-Flueh (SO)	18	64,902	55,863	0.61	11.5	190	8.77	24.43

† Cantonal and communal income tax rate for married couple with two children and income of CHF 80,000 and CHF 500,000 respectively.

Table 4.1 gives descriptive statistics of the household and community characteristics. From the total of 7,872 households that stayed within the choice set, 4/5 moved within the center community whereas only 1/5 moved to one of the 16 communities in the periphery. The latter were on average 20% richer than the ones remaining in the center. The tax rate of a typical two-child family with an income of CHF 80,000 is highest in the center community; this is more than 35% higher than in the neighboring community

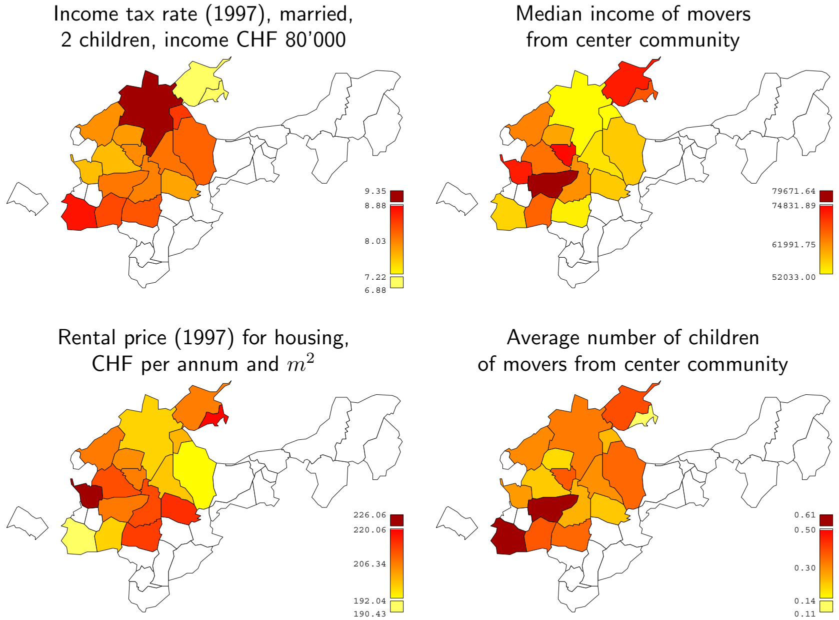

Figure 4.4: Characteristics of movers from the center community in 1997.

Riehen. The tax rate for an income of CHF 500,000 is about three times higher and the tax amount consequently 15 times as high, reflecting the strong progressivity of the different tax schemes. Figure 4.4 visualizes the association between the local tax level and the average income of households that moved in. The number of children of households in the center and in the periphery are very similar. However, there is substantial variation across the communities in the periphery. The bottom maps in Figure 4.4 show the local housing rent and the average number of children of the incomers. The center community is surrounded by a fringe of communities with higher rents. Families with more children tend to locate in communities with lower rents.

4.6 Results

The estimates of the random utility models with fixed effects are given in Table 4.2. Three specifications of the error term structure are reported: Column one shows the results for the conditional logit model, column 2 for the nested logit model and column three for the multinomial probit model with a spatial autoregressive process (SAR). All models are estimated with a full set of community-specific constants using the city of Basel as reference community.

The specification of the error term is discussed first. The nested logit model fits the data significantly better than the conditional logit model (likelihood ratio test statistic 5.4). The highly significant log-sum coefficient in the nested logit model clearly demonstrates the violation of the IIA assumption in the conditional logit model. The estimate $\hat{\lambda} = 0.53$ implies that the error terms across communities in the periphery are positively correlated. This means that households with an unobserved taste for a community in the periphery also prefer other communities in the periphery. The estimated spatial autocorrelation coefficient $\hat{\rho}$ in the multinomial probit model is not significantly different from zero. The spatial autocorrelation process therefore does not improve the probit model with independent error terms (log Likelihood = -7442.68), which is the analogue of the conditional logit model and also saddled with the IIA. The nested logit model is thus the preferred model. The following discussion relates to the results of the nested multinomial logit model.

While the sign and significance of the coefficients in multinomial response models are informative they cannot be directly interpreted as marginal effects. The significantly negative fixed effects of all communities in the periphery indicate that these communities are intrinsically less attractive to movers from the city of Basel. These estimates are not surprising as 4/5 of the movers decided to stay in the center. The fixed effects take account of locational factors such as housing prices, public goods provision, distance to the central business district, cultural activities and landscape but also a possible distaste for leaving the accustomed community.

The coefficient for $log(1 - tax)$ gives the effect of the tax rate on the indirect utility function for a household with an income of CHF 60,000, i.e. the average income. It is significantly positive at the 0.1% level and

Table 4.2: Multinomial response models with fixed effects.

	Condit. Logit			Nested Logit			Mult. Probit		
Slope Coefficients									
log(1-tax)†	15.40	(2.54)	***	14.55	(2.00)	***	8.06	(1.70)	***
log(1-tax)×[log(inc)-log(60k)]	11.28	(2.00)	***	11.04	(1.75)	***	7.59	(1.31)	***
log(rent)×[log(inc)-log(60k)]	0.70	(0.81)		0.56	(0.47)		0.26	(0.35)	
log(rent)× children	-1.32	(0.82)		-1.16	(0.56)	*	-0.80	(0.41)	~
Correlation Structure Coefficients									
Log-sum periphery (λ)				0.53	(0.08)	***			
Spatial autocorrelation (ρ)							0.46	(1.08)	
Fixed Effects									
Binningen (BL)	-4.14	(0.12)	***	-3.08	(0.20)	***	-2.69	(0.27)	***
Birsfelden (BL)	-4.54	(0.12)	***	-3.24	(0.23)	***	-2.86	(0.26)	***
Bottmingen (BL)	-5.45	(0.18)	***	-3.77	(0.31)	***	-3.31	(0.19)	***
Allschwil (BL)	-3.70	(0.11)	***	-2.84	(0.17)	***	-2.53	(0.17)	***
Münchenstein (BL)	-4.68	(0.13)	***	-3.35	(0.24)	***	-2.99	(0.14)	***
Oberwil (BL)	-4.90	(0.15)	***	-3.51	(0.26)	***	-3.12	(0.13)	***
Riehen (BS)	-3.52	(0.10)	***	-2.72	(0.15)	***	-2.47	(0.05)	***
Muttenz (BL)	-4.46	(0.12)	***	-3.23	(0.22)	***	-2.89	(0.14)	***
Bettingen (BS)	-6.93	(0.35)	***	-4.53	(0.45)	***	-3.87	(0.16)	***
Reinach (BL)	-4.17	(0.12)	***	-3.07	(0.20)	***	-2.79	(0.07)	***
Arlesheim (BL)	-5.21	(0.17)	***	-3.65	(0.29)	***	-3.25	(0.14)	***
Therwil (BL)	-5.35	(0.17)	***	-3.70	(0.30)	***	-3.31	(0.08)	***
Biel-Benken (BL)	-6.38	(0.26)	***	-4.29	(0.39)	***	-3.74	(0.14)	***
Aesch (BL)	-5.08	(0.16)	***	-3.53	(0.28)	***	-3.19	(0.08)	***
Ettingen (BL)	-5.97	(0.22)	***	-4.01	(0.35)	***	-3.55	(0.09)	***
Hofst.-Flueh (SO)	-6.06	(0.24)	***	-4.00	(0.37)	***	-3.53	(0.09)	***
Log likelihood	-7439.2			-7436.5			-7442.59		
Observations	7872			7872			7872		

Standard errors in brackets. ***, **, *, ~ significant at the 0.1%, 1%, 5% and 10% level.
†The coefficient gives the effect for a household with an income of CHF 60,000 as the interaction term $log(1-tax)\cdot[log(inc)-log(60,000)]$ becomes zero.

confirms that taxes have a negative effect on utility. The significantly positive coefficient of the interaction with $[log(inc)-log(60,000)]$ implies that the effect from taxes increases with income. For example, the effect from $log(1-tax)$ for a household with an income of CHF 500,000 is $14.55+11.04\cdot[log(500,000)-log(60,000)] = 37.96$. The quantitative impact of the tax rate will be explained using an example. Consider a household

with an income of CHF 500,000 that compares the city of Basel to the neighboring community of Riehen. The tax rate it faces is 26.4% in Basel and 21.8% in Riehen. The utility difference from this tax differential is $[log(1-0.218)-log(1-0.264)]\cdot 37.96 = 2.30$. Hence the negative fixed effect of Riehen (-2.72) is almost offset by its lower taxes. However, for a household with an income of CHF 80,000 the implied utility difference is only 0.5.

The coefficient of the local housing prices $log(rent)$ is not identified as its part is taken by the community constants. The interaction of the housing price with income can still be estimated but turns out to be insignificant. The housing rent also interacts with the number of children in the household. This effect is significantly negative at the 5% level. As one can sensibly suppose a negative effect of housing prices on utility, the negative sign of the interaction means that households with children are more concerned about housing prices than childless households.

How do these results correspond to the segregation hypotheses postulated in the theoretical model? The effect of the tax rate without interaction is identified only through the variation across households induced by the progressivity of the tax scheme. The significant positive sign of this coefficient establishes the income segregation from progressive taxation. Equation (4.9) depicts the conditions for segregation that are induced by mechanisms beyond the progressive tax scheme. The signs of the estimated coefficients satisfy this condition under the assumption that the unobserved price effect is negative. There is clear evidence that rich households prefer low-tax communities but to an even greater extent than is explained by the tax scheme.

The quantitative implications of the estimated nested logit model are revealed by inspecting its predictions. Given the attributes of the households and the communities, the model is able to predict the fraction of the households that move to a particular community j:

$$\frac{1}{N}\sum_{n=1}^{N} P_{nj}(\hat{\theta})\,. \tag{4.15}$$

Table 4.4 (column 3) shows the predicted migration to the communities. The actual values in the data set are given for comparison in column 1.

Table 4.3: Multinomial response models with community characteristics.

	Condit. Logit		Nested Logit		Mult. Probit	
Slope Coefficients						
log(1-tax)†	15.89 (2.26)	***	14.10 (1.98)	***	10.11 (1.25)	***
log(1-tax)×[log(inc)-log(60k)]	12.65 (1.91)	***	12.47 (1.68)	***	8.18 (1.15)	***
log(rent)†	-2.88 (0.77)	***	-1.84 (0.48)	***	-1.81 (0.35)	***
log(rent)×[log(inc)-log(60k)]	0.15 (0.67)		0.05 (0.37)		-0.02 (0.30)	
log(rent)×children	-0.96 (0.73)		-0.83 (0.46)	~	-0.60 (0.37)	
distance	-0.18 (0.01)	***	-0.09 (0.02)	***	-0.10 (0.01)	***
periphery	-3.47 (0.10)	***	-2.67 (0.20)	***	-2.29 (0.05)	***
Correlation Structure Coefficients						
Log-sum periphery (λ)			0.51 (0.10)	***		
Spatial autocorrelation (ρ)					0.76 (0.04)	***
Standard deviation rent ($\sigma_n u$)						
Log likelihood	-7778.1		-7774.5		-7759.6	
Observations	7872		7872		7872	

Standard errors in brackets. ***, **, *, ~ significant at the 0.1%, 1%, 5% and 10% level.
†The coefficient gives the effect for a household with an income of CHF 60,000 as the interaction term $log(1 - tax) \cdot [log(inc) - log(60,000)]$ becomes zero.

Note the perfect forecast of migration which is an artefact of the full set of community intercepts. However, these predictions will change in the policy experiments conducted in Section 4.6.1. The predicted segregation pattern is more informative. The average income of the households moving to community j is predicted as

$$\sum_{n=1}^{N} y_n P_{nj}(\hat{\theta}) \Big/ \sum_{n=1}^{N} P_{nj}(\hat{\theta}) \,. \tag{4.16}$$

Table 4.4 shows the predicted average income (column 4) of the migrants compared to the actual values (column 2). The top maps in Figure 4.5 visualize the actual income segregation and the segregation predicted by the nested logit model. As can be seen, the predicted pattern of income differences across communities is very similar to the observed pattern. This remarkably demonstrates the appropriateness of the econometric specification in equation (4.11).

The results of multinomial response models without fixed effects are given in Table 4.3. Remember that identifying locational factors is not the prime interest of this study and not much effort was spent on finding proxies

for community characteristics. The only additional variables are *distance* from the central business district and the now identified local housing prices $log(rent)$. A dummy variable for the periphery was also introduced to capture the high proportion of stayers. This dummy variable could as well have been labelled 'staying'.

The slope coefficients of the nested logit model are almost identical to the ones in the fixed effects model. As in the fixed effects model, the significant log-sum coefficient shows the violation of the IIA. The now identified housing price effect is significantly negative as was expected above. Not surprisingly, distance from the center has a significant negative impact.

The coefficient of spatial autocorrelation $\hat{\rho} = 0.73$ is now significantly positive. The predictions from the resulting multinomial probit model are very similar to the nested logit model. Note that the coefficients in the multinomial probit model are smaller than in the logit models because the model is scaled by the variance of the standard normal distribution $\sigma_\varepsilon = 1$ rather than that of the extreme value distribution $\sigma_\varepsilon = 1.7$.

4.6.1 Policy Experiments

An important feature of the estimated models is that they can be used to simulate the aggregate effects from changes in policy variables. Given the attributes of the households and the communities after the implementation of the experiment, the models can predict the number of migrants and their average income according to equations (4.15) and (4.16). Two experiments are performed.

Experiment 1 (Tax increase in the center)
The center community increases its income tax rate by a factor of 1.1 for all household types.

Table 4.4 shows the predicted effects of the two experiments based on the estimated nested logit model with fixed effects. Experiment 1 means that the tax rate of a two-child family with an income of CHF 80,000 rises by almost one percentage point from 9.4% to 10.3% in the center community. The tax increase is 2.6% for an income of CHF 500,000. As a result, fewer households would choose the center community. This can be seen in column 5 of Table 4.4. The center community would lose 3.7% of the moving

Table 4.4: Model predictions and results of policy experiments.

	Sample Values		Model		Experiment 1		Experiment 2	
	share mover	mean income	share mover	mean income	share mover	mean income	share mover	mean income
Whole area	1	61,612	1	61,612	1	61,612	1	61,612
City of Basel (BS)	0.809	59,334	0.809	59,231	0.772	56,708	0.817	59,619
Periphery	0.191	71,271	0.191	71,710	0.228	78,233	0.183	70,522
- Binningen (BL)	0.021	73,405	0.021	73,362	0.025	80,052	0.022	75,375
- Birsfelden (BL)	0.012	52,351	0.012	63,702	0.015	68,800	0.013	65,316
- Bottmingen (BL)	0.005	76,376	0.005	71,473	0.007	77,729	0.006	73,370
- Allschwil (BL)	0.032	69,302	0.032	73,655	0.038	80,444	0.034	75,695
- Münchenstein (BL)	0.012	58,962	0.012	69,193	0.014	75,276	0.012	71,056
- Oberwil (BL)	0.010	77,048	0.010	80,627	0.012	88,895	0.011	83,093
- Riehen (BS)	0.036	83,950	0.036	75,489	0.042	82,913	0.018	52,069
- Muttenz (BL)	0.014	63,333	0.014	66,291	0.017	71,905	0.015	68,047
- Bettingen (BS)	0.001	69,978	0.001	86,834	0.001	98,430	0.001	53,220
- Reinach (BL)	0.019	72,242	0.019	69,510	0.023	75,100	0.020	71,221
- Arlesheim (BL)	0.007	57,601	0.007	74,243	0.009	80,690	0.008	76,181
- Therwil (BL)	0.006	91,735	0.006	67,810	0.007	73,198	0.006	69,477
- Biel-Benken (BL)	0.002	88,610	0.002	84,549	0.003	93,477	0.002	87,255
- Aesch (BL)	0.007	62,968	0.007	65,115	0.009	69,965	0.008	66,635
- Ettingen (BL)	0.003	61,541	0.003	62,628	0.004	67,365	0.003	64,156
- Hofst.-Flueh (SO)	0.002	64,902	0.002	55,773	0.003	60,061	0.002	57,331

Notes: Predictions and results from experiments (see text) using the estimated nested logit model with fixed effects.

population. The change of the spatial distributions of incomes provided in column 6 is particularly interesting. The average income would fall by CHF 2,500 in the center community and increase by CHF 6,500 in the periphery, ranging from CHF 4,300 in Hofstetten-Flueh to 11,600 in Bettingen. These effects are depicted in the bottom-left map in Figure 4.5.

Experiment 2 (Tax harmonization in one canton)

The Canton of Basel-Stadt decides that its two communities in the periphery (Riehen and Bettingen) will employ the same tax scheme and rate as its capital, the city of Basel.

Experiment 2 simulates the choice of moving households if the Canton of Basel-Stadt harmonized its taxes and applied the city's tax rates in its

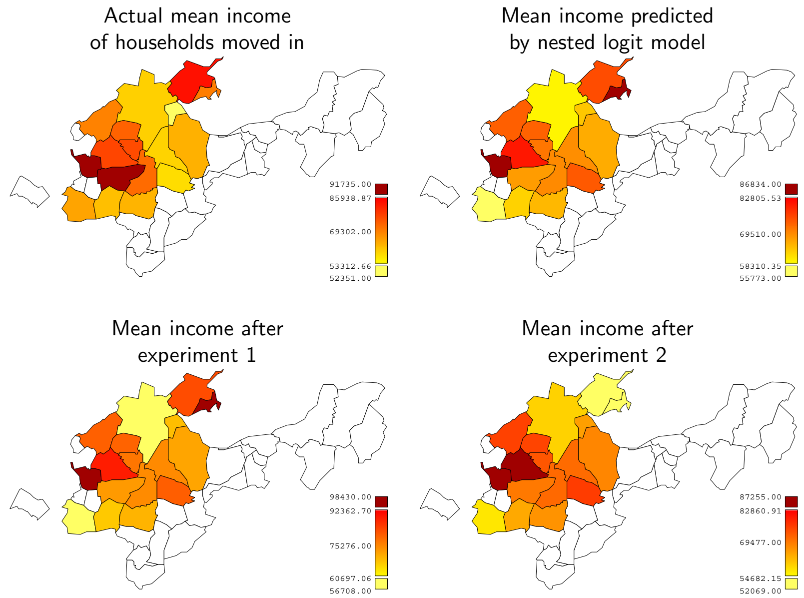

Figure 4.5: Model predictions and results of policy experiments.

other communities. The fairly dramatic effects are given in columns 7 and 8 in Table 4.4 and in the bottom-right map in Figure 4.5. The two peripheral communities Riehen and Bettingen would lose half of the new households. More interestingly, the average income of households that moved to Riehen and Bettingen would fall by CHF 23,400 and CHF 33,600, respectively. The average income of migrating households would increase by a meager CHF 400 in the city of Basel but by almost CHF 2000 in the other communities in the periphery.

It is important to keep in mind that the simulated effects neglect three potential sources of bias. Firstly, the estimated model neglects the possible migration to the center from households currently living in the periphery. Secondly, households that did not move under the actual situation may decide to move after the policy change and actually moving households may remain in their homes. Thirdly, the estimated model ignores that community characteristics such as housing prices and public goods provision would adjust to policy changes. However, this reaction is likely to be a long-run effect while moving households immediately adapt their community choice.

4.7 Conclusions and Outlook

Theoretical models of urban community systems postulate endogenous segregation of the population by incomes. The location choice of households with differing incomes is supposed to be cause and consequence of the local tax rate differentials. This chapter empirically assesses the second causal connection by studying the community choice of households in a Swiss metropolitan area.

The estimation results show that rich households are substantially and significantly more likely to move to low-tax communities than poor households. This self-sorting of the migrating population perpetuates the existing income differentials across communities. The higher esteem of rich households for low taxes is partly explained by the progressivity of the local income tax. However, there is evidence that rich households prefer low-tax communities to a greater extent than is explained by the tax scheme.

The observed sorting of the population by incomes could possibly be influenced by factors not considered in this study. A promising alternative explanation is 'social interaction': If rich households preferred to live near other rich households a similar segregation pattern would emerge. Such neighborhood interaction was found by Ioannides and Zabel (2002) and Bayer, McMillan and Rueben (2002). Unfortunately, the present data do not allow to discriminate between the two explanations because the average local income levels and the local tax rates are almost multicollinear. A possible way to overcome this problem lies in inspecting the location choice below community level. The differences of average incomes across districts within the same community can be exploited to identify both the effects of the tax rates and of the neighborhood characteristics.

A possible further extension of this study is the collection of more information on potential locational factors, such as school quality. This additional information can be used to study more interaction effects with household characteristics. School quality for example is most likely an important

locational factor for families with children, but not for single households. However, as argued in the beginning of Section 4.3, differences in the provision of public goods do not seem essential in Switzerland.

4.A Appendix: Data

The data used in the empirical investigation were made available by the following institutions:

Household data	Statistical Office of the Canton of Basel-Stadt, merged data from the Cantonal Tax Administration and the Residents Registration Office.
Tax schemes	Swiss Federal Tax Administration, Steuerbelastung in der Schweiz, Natürliche Personen nach Gemeinden 1997, Neuchâtel: Swiss Federal Statistical Office.
Housing prices	Wüest und Partner, Zurich.
Income Distribution	Swiss Federal Tax Administration, Direkte Bundessteuer 1993/94 - Gemeinden.
Commuter	Swiss Federal Statistical Office, Census 1990.

Notes on the construction of the variables:

Income (uses household data): The information on the household income is based on the tax assessment. Unmarried adult household members and children with their own income are assessed individually. The income of all individually assessed household members is added up. The income in the raw data is income before tax and deductions for children and spouse but after social security contributions and further deductions. The study uses (hypothetical) gross income which was calculated without considering further individual deductions.

Children (uses household data): Number of children that allow for tax deductions.

Tax rate (uses household and tax scheme data): The tables from the Swiss Federal Tax Administration report the totalled cantonal and communal tax rates for different household types (single household, married couple without children and married couple with two children) and for selected gross incomes. The tax rate for households with income between the reported income classes and for household types not listed were interpolated. The tax rates for household members with individual tax assessment were first calculated individually. The tax rate of the household is calculated from the totalled individual tax amounts.

Rent (uses housing price data): Wüest und Partner collected all rents for flats offered in newspapers and in the internet in 1997. Missing information on exact flat sizes was inferred from the information given in the advertisements.

Distance: Distance between the geographical centers of the communities. The center was taken as the middle of the maximal east-west and north-south extensions.

Bibliography

Anselin, Luc and Raymond J.G.M Florax (1995), Eds., *New Directions in Spatial Econometrics*, Berlin: Springer.

Bayer, Patrick, Robert McMillan, and Kim Rueben (2003), An Equlibrium Model of Sorting in an Urban Housing Market: A Study of the Causes and Consequences of Residential Segregation, CES Research Paper 03-01, U.S. Bureau of the Census.

Berman, Abraham and Robert J. Plemmons (1994), *Nonnegative Matrices in the Mathematical Sciences*, Philadelphia: Siam.

Berry, Steven, James Levinsohn and Ariel Pakes (1995), Automobile Prices in Market Equilibrium, *Econometrica*, 63(4), 841-890.

Bhat, Chandra R. (2001), Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model, *Transportation Research B*, 35, 677-693.

Bolduc, Denis (1992), Generalized Autoregressive Errors: The Multinomial Probit Model, *Transportation Research B*, 26, 155-70.

Bolduc, Denis, Bernard Fortin and Marc-Andre Fournier (1996), The Impact of Incentives Policies on the Practice Location of Doctors: A Multinomial Probit Analysis, *Journal of Labor Economics*, 14, 703-32.

Bolduc, Denis, Bernard Fortin and Stephen Gordon (1997), Multinomial Probit Estimation of Spatially Interdependent Choices: An Empirical Comparison of Two New Techniques, *International Regional Science Review*, 20, 77-101.

Börsch-Supan, Axel (1990), On the Compatibility of Nested Logit Models with Utility Maximisation, *Journal of Econometrics*, 43, 373-388.

Börsch-Supan, Axel and Vassilis A. Hajivassiliou (1993), Smooth Unbiased Multivariate Probabilities Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models, *Journal of Econometrics*, 58, 347-368.

Calabrese, Stephen A. (1990), Local Redistribution financed by Income Tax, Carnegie Mellon University Working Paper.

Deateon, Angus and John Muellbauer (1980), *Economics and Consumer Behavior*, Cambridge: Cambridge University Press.

Ellickson, Bryan (1971), Jurisdictional Fragmentation and Residential Choice, *American Economic Review*, 61(2), 334-339

Epple, Dennis, Radu Filimon and Thomas Romer (1984), Equilibrium among Local Jurisdictions: Towards an Integrated Treatment of Voting and Residential Choice, *Journal of Public Economics*, 24, 281-308.

Epple, Dennis, Radu Filimon and Thomas Romer (1993), Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes, *Regional Science and Urban Economics*, 23, 585-610.

Epple, Dennis and Glenn J. Platt (1998), Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income, *Journal of Urban Economics*, 43, 23-51.

Epple, Dennis and Thomas Romer (1991), Mobility and Redistribution, *Journal of Political Economy*, 99(4),828-858.

Epple, Dennis, Thomas Romer and Holger Sieg (2001), Interjurisdictional Sorting and Majority Rule: An Empirical Analysis, *Econometrica*, 69(6), 1437-1465.

Epple, Dennis and Holger Sieg (1999), Estimating Equilibrium Models of Local Jurisdictions, *Journal of Political Economy*, 107, 645-681.

Feld, Lars P. (1999), *Steuerwettbewerb und seine Auswirkungen auf Allokation und Distribution: Eine empirische Analyse für die Schweiz*, Dissertation Nr. 2222, Universität St. Gallen.

Feld, Lars P. and Gebhard Kirchgässner (2001), Income Tax Competition at the State and Local Level in Switzerland, *Regional Science and Urban Economics*, 31, 181-213.

Friedman, Joseph (1981), A Conditional Logit Model of the Role of Local Public Services in Residential Choice, *Urban Studies*, 18, 347-358.

Geweke, John, Michael Keane and David Runkle (1994), Alternative Computational Approaches to Inference in the Multinomial Probit Model, *Review of Economics and Statistics*, 1994, 609-632.

Goodspeed, Timothy J. (1986), *Local Income Taxation: Equilibrium, Social Optimality and Welfare Loss in a General Equlibrium Metropolitan Model*, Ph.D. Dissertation, University of Maryland.

Goodspeed, Timothy J. (1989), A Re-Examination of the Use of Ability to Pay Taxes by Local Governments, *Journal of Public Economics*, 38, 319-342.

Guesnerie, R. and J.-J. Laffont (1984), A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of Self-Managed Firm, *Journal of Public Economics*, 25, 329-369.

Hansen, Nico A. and Anke S. Kessler (2001a), The Political Geography of Tax H(e)avens and Tax Hells, *American Economic Review*, 91(4), 1103-1115.

Hansen, Nico A. and Anke S. Kessler (2001b), (Non-)Existence of Equilibria in Multi-Community Models, *Journal of Urban Economics*, 50, 418-435.

Kessler, Anke S. and Christoph Lülfesmann (1999), Tiebout and Redistribution in a Model of Residential and Political Choice, unpublished working paper, University of Bonn.

Hajivassiliou, Vassilis A. and Paul A. Ruud (1994), Classical Estimation Methods for LDV Models using Simulation. In: R. F. Engle and D.L. McFadden, Eds., *Handbook of Econometrics IV*, Amsterdam: Elsevier.

Heiss, Florian (2002), Structural choice analysis with nested logit models, *The Stata Journal*, 2(3), 227-252.

Hensher, David A. and William H. Greene (2002), Specification and Estimation of Nested Logit Models, *Transportation Research B*, 36(1), 1-18.

Hilber, Christian (1998), *Auswirkungen staatlicher Massnahmen auf die*

Bodenpreise. Eine theoretische und empirische Analyse der Kapitalisierung, Chur: Rüegger.

Ioannides, Yannis M. and Jeffrey E. Zabel (2002), Interactions, Neighborhood Selection and Housing Demand, Department of Economics Discussion Paper 2002-08, Tufts University.

Judd, Kenneth L. (1998), *Numerical Methods in Economics*, Cambridge: MIT Press.

McFadden, Daniel (1974), Conditional Logit Analysis of Qualitative Choice Behavior. In: P. Zarembka, Ed., *Frontiers in Econometrics*, New York: Academic Press.

McFadden, Daniel (1978), Modelling the Choice of Residential Location. In: A. Karlqvist et al., Eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland.

McFadden, Daniel (1984), Econometric Analysis of Qualitative Response Models. In: Zvi Griliches and Michael D. Intriligator, Eds., *Handbook of Econometrics II*, Amsterdam: Elsevier.

McFadden, Daniel (2001), Economic Choices, *American Economic Review*, 91, 351-378.

Nechyba, Thomas J. and, Robert P. Strauss (1998), Community Choice and Local Public Services: A Discrete Choice Approach, *Regional Science and Urban Economics*, 28, 51-73.

Oates, Wallace E. (1969), The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis, *Journal of Political Economy*, 77, 951-971.

Oates, Wallace E. (1972), *Fiscal Federalism*, New York: Harcourt Brace Jovanovich.

Rhode, Paul W. and Koleman S. Strumpf (forthcoming), Assessing the Importance of Tiebout Sorting: Local Heterogeneity form 1850 to 1990, *American Economic Review*.

Rose-Ackerman, Susan (1979), Market Models of Local Government: Exit, Voting, and the Land Market, *Journal of Urban Economics*, 6, 319-337.

Ross, Stephen and John Yinger (1999), Sorting and Voting: A Review of the Literature on Urban Public Finance. In: Paul Ceshire and Edwin S. Mills, Eds., *Handbook of Regional and Urban Economics*, Vol. 3, Amsterdam: North-Holland.

Tiebout, Charles M. (1956), A Pure Theory of Local Expenditures, *Journal of Political Economy*, 64(5), 416-424.

Train, Kenneth E. (2003), *Discrete Choice Methods with Simulation*, Cambridge: Cambridge University Press.

Westhoff, Frank (1977), Existence of Equilibria in Economies with a Local Public Good, *Journal of Economic Theory*, 14, 84-112.

Young, Peyton H. (1990), Progressive Taxation and Equal Sacrifice, *American Economic Review*, 80(1), 253-266.