Vermögen in Österreich

Bericht zum Forschungsprojekt „Reichtum im Wandel“[*]

Juli 2013

Paul Eckerstorfer
Institut für Volkswirtschaftslehre
Johannes Kepler Universität Linz
Email: paul.eckerstorfer@jku.at

Johannes Halak
Institut für Philosophie und Wissenschaftstheorie
Johannes Kepler Universität Linz
Email: hannes.halak@jku.at

Jakob Kapeller
Institut für Philosophie und Wissenschaftstheorie
Johannes Kepler Universität Linz
Email: jakob.kapeller@jku.at

Bernhard Schütz
Institut für Volkswirtschaftslehre
Johannes Kepler Universität Linz
Email: bernhard.schuetz@jku.at

Florian Springholz
Institut für Philosophie und Wissenschaftstheorie
Johannes Kepler Universität Linz
Email: florian.springholz@jku.at

Rafael Wildauer
Institut für Philosophie und Wissenschaftstheorie
Johannes Kepler Universität Linz
Email: rafael.wildauer@jku.at

JOHANNES KEPLER UNIVERSITÄT LINZ | JKU

[*] Gefördert aus Mitteln der Arbeiterkammern Wien und Oberösterreich.

Executive Summary

Der Household Finance and Consumption Survey (HFCS) der europäischen Zentralbank (EZB) stellt die erste umfassende Erhebung zu Vermögen privater Haushalte in 15 Ländern der Eurozone, darunter auch Österreich, dar. Somit ermöglicht der HFCS erstmalig eine genaue Analyse der österreichischen Vermögensbestände sowie der Vermögensverteilung. Trotz akribischer Erhebung der Daten durch die Österreichische Nationalbank (OeNB) und das Institut für empirische Sozialforschung (IFES) besteht bei dieser Erhebung das Problem der fehlenden oder unzureichenden Erfassung der obersten Vermögensbestände, die in den Händen einiger weniger Haushalte konzentriert sind. Damit geht eine systematische Unterschätzung des Gesamtvermögens privater Haushalte in Österreich sowie eine Verzerrung der tatsächlichen Vermögensverteilung einher. Um diese Verzerrung zu kompensieren empfiehlt die einschlägige wissenschaftliche Literatur die Verwendung der Pareto-Verteilung. Bei dieser Methode wird unter Zuhilfenahme statistischer Tests postuliert, dass sich der oberste Rand der Vermögensverteilung durch eben jene Pareto-Verteilung näherungsweise darstellen lässt. In dem solcherart korrigierten Datensatz steigt das Gesamtvermögen der privaten Haushalte von etwa 1.000 Mrd. Euro auf 1.249 Mrd. Euro an. Besonders stark wirkt sich die Korrektur der Nicht- und Untererfassung auf den Vermögensbestand des reichsten Prozents aller Haushalte aus. Dieser steigt von durchschnittlich 6,4 Millionen Euro um 98,6% auf 12,7 Millionen Euro. Daraus ergibt sich unter anderem, dass die reichsten 10% der ÖsterreicherInnen nicht 61% (HFCS) sondern 69% des Gesamtvermögens besitzen.

Inhaltsverzeichnis

1. Einleitung 5

2. Vermögen in Österreich: Eine erste deskriptive Analyse 7

2.1. Household Finance and Consumption Survey (HFCS) 7

2.2. Vermögen in Österreich: Bisherige Ergebnisse im Überblick 9

2.3. HFCS-Verteilungsstatistiken: Eine deskriptive Analyse 10

3. Eine Schätzung der Verteilung des österreichischen Privatvermögens 14

3.1. Statistische Aspekte einer Theorie der Vermögensverteilung 15

3.2. Alternative Strategien zur Quantifizierung von Top-Vermögen 18

3.3. Schätzverfahren und Datenbearbeitung 21

3.4. Die Vermögensverteilung in Österreich unter Berücksichtigung der Datenkorrektur 26

4. Resümee 29

5. Literatur 31

Anhang

Anhang I: Perzentilliste auf Basis der HFCS-Daten 34

Anhang II: Schätzung der Pareto-Verteilung: Berechnungen (Mathematica-Code) ... 35

Anhang III: Perzentilliste auf Basis der korrgierten Daten 39

Anhang IV: Samplekorrektur: Berechnungen (Mathematica-Code) 40

Abbildungsverzeichnis

Abbildung 1: Die Vermögensverteilung in Österreich nach Vermögensklassen auf Basis der HFCS-Daten (Replikation einer OeNB-Graphik; Original in: OeNB 2012, 41) 11
Abbildung 2: Kumulative Verteilungsfunktion der österreichischen Privatvermögen auf Basis der HFCS-Daten 12
Abbildung 3: Lorenz-Kurve auf Basis der HFCS-Daten 13
Abbildung 4: Darstellung von relativen Vermögensanteilen mittels der HFCS-Daten 14
Abbildung 5: Veranschaulichung der Methode zur Daten-Korrektur 21
Abbildung 6: Geschätzte Alpha-Parameter und dazugehörige p-Werte (nach Cramer-von-Mises) für die obersten 30 Perzentilgrenzen (unter Berücksichtigung aller Imputationen) 23
Abbildung 7: Die Vermögensverteilung in Österreich nach Vermögensklassen auf Basis der *modifizierten* HFCS-Daten (Replikation von Abbildung 1; Unterschiede in Orange) 27
Abbildung 8: Lorenz-Kurve mit modifizierten HFCS-Daten 28
Abbildung 9: Darstellung von relativen Vermögensanteilen mittels der korrigierten HFCS-Daten 29

Tabellenverzeichnis

Tabelle 1: Vermögensverteilung der obersten 5 Perzentile 12
Tabelle 2: Nettovermögen und geschätzte Pareto-Alphas am Schwellenwert des 78. Perzentils 24
Tabelle 3: Die Vermögen der reichsten fünf Perzentile auf Basis der modifizierten HFCS-Daten 28

1. Einleitung

Gesellschaftliche Stabilität und privater Vermögensaufbau stehen in einem ambivalenten Zusammenhang. Die historische Perspektive zeigt, dass gesellschaftliche Stabilität und etablierte Institutionen eine zentrale Voraussetzung für den Aufbau von Vermögenswerten darstellen (Atkinson et al. 2011, Borgherhoff/Mulder et al. 2009). Allerdings führt der Aufbau privater Vermögen zu sich selbst verstärkenden Effekten: Jene, die bereits Vermögen besitzen, erhalten im Schnitt auch (absolut wie relativ) größere Vermögenzuwächse. Dies führt zu einer sukzessiv ansteigenden Vermögenskonzentration und dem Aufbau von Schulden, die den angehäuften Finanzvermögen gegenüberstehen.

Eine zunehmende Vermögenskonzentration kann also ihrerseits die gesellschaftliche Stabilität aus sozialer wie ökonomischer Perspektive untergraben (Guttmann/Plihon 2010, Stiglitz 2012). In diesem Sinne ist die wissenschaftliche Auseinandersetzung mit privaten Vermögen und deren Konzentration von hoher gesellschaftlicher und ökonomischer Bedeutung.

Den hier unterstellten Zusammenhang zwischen sozialem Zusammenhalt, gesellschaftlicher Stabilität und ökonomischen Verteilungsergebnissen greifen auch Wilkinson und Pickett (2007) auf. Ihre Studie, deren Ergebnisse unter anderem in dem Buch „The Spirit Level“ (2009) publiziert wurden, zeigt, dass ein höheres Maß an Gleichheit in einer Gesellschaft zahlreiche gesellschaftlich erwünschte Effekte mit sich bringt. Diese werden ihrerseits als Indikatoren für die gesellschaftliche Zufriedenheit interpretiert. Die StudienautorInnen zeigen empirisch, dass das Niveau gesellschaftlicher Ungleichheit in einem direkten Zusammenhang mit gesellschaftlichen Problemstellungen wie Fettleibigkeit, Teenager-Schwangerschaften, Lebenserwartung, mentalen Krankheiten, Selbstmordraten, Fremdenfeindlichkeit, Drogenkonsum, Bildungsperformance und Inhaftierungsraten steht (vgl. Wilkinson/Pickett 2007).

Mit der Verbesserung computergestützter Auswertungsmethoden und allgemein durch die steigende Verfügbarkeit von Erhebungsdaten lässt sich die Ungleichheit in der Verteilung von Vermögen zusehends besser erfassen. Zusätzlich erleichtert die ex-

ante Harmonisierung von Erhebungen, wie sie auch dem von der EZB koordinierten und den einzelnen europäischen Nationalbanken durchgeführten *Household Finance and Consumption Survey* (HFCS) zu Grunde liegen, die Vergleichbarkeit der Daten zwischen den Ländern. Für den Fall Österreichs, aber auch für viele andere Länder der Eurozone, stellt der HFCS die bis dato bei Weitem umfassendste haushaltsbezogene Erhebung zur Vermögenssituation und insbesondere zu den disaggregierten Vermögensbestandteilen (Finanzvermögen, Sachvermögen, Schulden etc.) dar und erlaubt damit einen bislang nicht möglichen Einblick in die Vermögenssituation österreichischer Privathaushalte.

Allerdings sind derartige Erhebungen mit zahlreichen Problemstellungen konfrontiert, die antizipiert werden müssen, um zu einer möglichst realitätsgetreuen Einschätzung zu gelangen: sie enthalten falsche Angaben, zahlreiche Antwortverweigerungen und repräsentieren die Vermögensverteilung nicht vollständig, da die oberste Spitze der VermögensinhaberInnen zumeist gar nicht in der Befragung auftaucht (Hoeller et al. 2012; Avery/Elliehausen/Kennickell 1986).

Während die meisten dieser Nachteile im Zuge des Survey-Designs des HFCS Berücksichtigung finden und daher im Rahmen der Durchführung der Befragung antizipiert werden, korrigiert das Design des HFCS nicht für die fehlende Repräsentativität bei den größten Vermögen. Dies stellt eine beachtenswerte Einschränkung der Anwendbarkeit der HFCS-Erhebung dar, die für eine Reihe öffentlich diskutierter Aspekte der österreichischen Vermögenssituation – etwa Fragen nach Vermögenskonzentration und Vermögensbesteuerung – relevant ist. Schließlich besitzt die betreffende Personengruppe einen hohen Anteil des Gesamtvermögens und ist für eine seriöse Aufarbeitung der Vermögenssituation daher von besonderer Bedeutung.

Zum Ausgleich dieser Verzerrung der HFCS-Daten soll in dieser Arbeit eine Verteilungsschätzung für die reichsten Haushalte herangezogen werden, die auf den HFCS-Daten selbst beruht. Im Erstellungsprozess dieser Studie wurden dabei unterschiedliche technische Wege durchdacht. Eine Möglichkeit besteht in der Schätzung einer log-normalisierten Verteilung, die jedoch wenig geeignet erscheint, besonders reiche Haushalte am oberen Ende der Verteilung abzubilden. In der

Literatur weit verbreitete Herangehensweisen sind hingegen die Verwendung sowohl der Pareto-, als auch der Dagum- und der Singh-Maddala-Verteilung. Für den vorliegenden Fall zeigt sich, dass insbesondere für die Abbildung besonders reicher Haushalte die Pareto-Verteilung eine gangbare und anwendungsorientierte Lösung bietet. Der Grad der Verzerrung der HFCS Daten, welcher durch die Unter- bzw. Nichtrepräsentation besonders reicher Haushalte entsteht, wird im Weiteren anhand gängiger Verteilungsstatistiken beleuchtet.

Der Forschungsbericht orientiert sich damit an folgendem Aufbau: In Kapitel 2 wird der Datensatz des HFCS vorgestellt, mit dessen Hilfe die Vermögensverteilung in Österreich deskriptiv dargestellt wird. Das dritte Kapitel nimmt eine Korrektur der Verteilungsdaten anhand der obig skizzierten Methoden vor. In diesem Kontext wird auch die dieser Strategie zu Grunde liegende theoretische Argumentation genauer dargestellt. Die hier entwickelte korrigierte Verteilungsstatistik wird in Folge mit den Originaldaten verglichen.

2. Vermögen in Österreich: Eine erste deskriptive Analyse

Das Ziel dieses Kapitels ist eine Darstellung der im Forschungsbericht verwendeten Daten aus dem Household Finance and Consumption Survey (HFCS). Ausgewählte deskriptive Statistiken sollen hierbei einen Überblick über die Vermögenssituation in Österreich vermitteln, wie sie von den HFCS-Originaldaten dargestellt wird.

2.1. Household Finance and Consumption Survey (HFCS)

Mit dem *„Household Finance and Consumption Survey“* (HFCS) steht erstmals eine umfassende Erhebung zu Sachvermögen, Finanzvermögen, Verbindlichkeiten und Ausgaben privater Haushalte in 15 Ländern der Eurozone (mit Ausnahme von Irland und Estland) zur Verfügung. Insbesondere die prinzipielle Vergleichbarkeit der Daten, die durch eine ex-ante Harmonisierung der Erhebung in den teilnehmenden Ländern sichergestellt wurde, lässt wertvolle Vergleiche zwischen Ländern der Eurozone zu. Die Erhebung der österreichischen Daten wurde von der Österreichischen Nationalbank (OeNB) gemeinsam mit dem Institut für empirische Sozialforschung GmbH (IFES) durchgeführt. Der folgende Abschnitt soll das Design

der Erhebung sowie ihre wichtigsten Bestandteile und Besonderheiten darstellen. Eine umfassende Dokumentation des HFCS findet sich auch online unter: http://www.hfcs.at. Eine genaue Darstellung des hier nur grob umrissenen Survey-Designs findet sich außerdem im Addendum zu den methodischen Grundlagen des HFCS (vgl. Albacete et al. 2012)

2.1.1 Stichprobe und Erhebungseinheit

Im Rahmen des HFCS wurden in Österreich 4.436 Haushalte in die Bruttostichprobe aufgenommen, wobei 2.380 Haushalte erfolgreich interviewt wurden und somit die Nettostichprobe bilden. Die Erhebungseinheit des HFCS ist ein Haushalt, wobei ein Haushalt als Person oder gemeinsam wirtschaftende Personengemeinschaft definiert ist. Im Rahmen der Durchführung der Erhebung wurde einE KompetenzträgerIn als Auskunftspersonen im Haushalt bestimmt. Der Zeitraum der Durchführung war September 2010 bis Mai 2011. Zur Auswahl der Stichprobe wurde das Verfahren einer mehrfach geschichteten Zufallsstichprobe (*Stratified Multistage Cluster Random Sampling*) angewandt. Damit wird jedem Element der Grundgesamtheit (d.h. allen Haushalten in Österreich) eine positive Wahrscheinlichkeit zugeordnet, um in die Stichprobe zu gelangen, wobei die Stratifizierung anhand der NUTS-3 Regionen in Österreich und nach Gemeindegrößenklassen sicherstellen soll, dass Haushalte aus verschiedenen Regionen proportional in der Stichprobe wiederzufinden sind.

2.1.2 Editierungsmaßnahmen und Konsistenzprüfungen

Die Erhebung wurde mithilfe einer computergestützten Befragung (*Computer Assisted Personal Interviews*) durchgeführt und erlaubt so schon während der Befragung das Ausschließen von Inkonsistenzen bzw. logisch unmöglichen Antworten. Auch nach Durchführung der Befragung wurden die Daten stichprobenartig einer expertenbasierten Analyse zugeführt, um die Datenkonsistenz durch nachträgliche Recherchearbeiten weiter zu erhöhen.

2.1.3 Multiple Imputationen

Typisch für alle freiwilligen Befragungen ist es, dass einige ErhebungsteilnehmerInnen gewisse Antworten nicht geben können oder nicht geben wollen (*item non response*). Vor allem bei Datenerhebungen zu „sensiblen“ Bereichen

wie Vermögen, Einkommen oder Schulden kann der Anteil an Antwortverweigerungen zu einer Verzerrung der Datenlage führen, der jedoch mittels der Methode der multiplen Imputation entgegengewirkt werden kann. Fehlende Werte können mittels multipler Imputation jeweils durch mehrere geschätzte Werte ersetzt werden, was dazu beiträgt, die ursprüngliche Korrelationsstruktur des Datensatzes zu bewahren, anstatt wie sonst üblich die gesamte Beobachtung aufgrund fehlender Angaben zu löschen. Insgesamt wurde dieser Prozess für jede Beobachtung fünfmal wiederholt. Der daraus resultierende Datensatz, der letztendlich von Seiten der OeNB zur Verfügung gestellt wurde, besteht damit aus fünf imputierten Samples - den sogenannten *Implicates*. Systematisch verzerrte Angaben (beispielsweise eine erhöhte Anzahl von Falschangaben in gewissen Segmenten der Vermögensverteilung) und das Problem der Nichtteilnahme besonders vermögender Haushalte (ein Problem der Abdeckung, engl.: *coverage)* können hingegen durch die Strategie der multiplen Imputation nicht kompensiert werden.

2.1.4 Gewichtungen

Zur Verringerung der Stichprobenvarianz und zur Anpassung der Stichprobe an die Zielpopulation wurden die Datensätze gewichtet. Die Gewichtung erfolgt aufgrund der ungleichen Wahrscheinlichkeiten, in die Stichprobe zu gelangen (*Unequal Probability Sampling Bias*), möglichen Verzerrungen durch die Unvollständigkeit der Auswahlpopulation (*Frame Bias*) und aufgrund von Antwortverweigerungen (*Non-Response Bias).* Die letztendlichen Gewichte ergeben sich damit aus den Design-Gewichten (w_{Di}), welche die ungleiche Selektionswahrscheinlichkeit korrigieren sollen, den Non-Response Gewichten (w_{NRi}) und den Post-Stratifizierungsgewichten (w_{PSi}), die einen irrtümlichen Ausschluss von Haushalten (z.B. falsche Postadresse) korrigieren sollen.

2.2. Vermögen in Österreich: Bisherige Ergebnisse im Überblick

Bisherige Studien zur österreichischen Einkommens- und Vermögensverteilung sind mit der Problematik einer unzureichenden Datenlage konfrontiert und erfassen meist nur einzelne Vermögensbestandteile wie das Finanzvermögen. Damit können sie zwar einen Eindruck der Verteilung vermitteln, ohne diese jedoch hinreichend abbilden zu können.

Beer et al. (2006) kommen auf Basis von Erhebungsdaten der Österreichischen Nationalbank (OeNB) aus dem Jahr 2004 zum Schluss, dass das durchschnittliche Nettogeldvermögen österreichischer Haushalte 51.790 EUR betrug (Median: 21.855 EUR) und verweisen unter anderem auf den starken Zusammenhang zwischen (ungleicher) Einkommensverteilung und (ungleicher) Geldvermögensverteilung in Österreich (vgl. Beer et al. 2006: 103). Die Analyse von Hahn und Magerl (2006) versucht als weiteren Ansatz die Daten der gesamtwirtschaftlichen Vermögensrechnung zu disagreggieren und daraus die Verteilung der Teilaggregate zu schätzen (vgl. Hahn/Magerl 2006). Das Ergebnis dieser Studie ist konsistent mit jenen von Beer et al. (2006).

Auf Basis von Lohnsteuerstatistiken und Beitragsstatistiken des Hauptverbandes der österreichischen Sozialversicherungsträger kommen Guger und Marterbauer (2007) im Rahmen einer WIFO-Studie zum Schluss, *„dass die Schere zwischen niedrigen und hohen Einkommen groß und in den letzten Jahrzehnten aufgegangen ist“* (Guger/Marterbauer 2007:1), bemängeln jedoch, dass die *„Verteilungsanalyse durch Datenmangel erheblich beeinträchtigt“* (Guger/Marterbauer 2007:23) werde. Andreasch, Fessler und Schürz (2009) errechnen unter Verwendung von Mikrodaten des Firmenbuchs ein Beteiligungsvolumen privater Haushalte an Aktien von rund 22,3 Mrd. EUR. Ein Abgleich mit der OeNB-Geldvermögenserhebung (2004) deutet insgesamt eine extreme Konzentration der Beteiligung an Kapitalgesellschaften unter besonders reichen Haushalten an (vgl. Andreasch/Fessler/Schürz 2009: 79f). Unter Bezugnahme auf die OeNB Immobilienvermögensbefragung (2008) schätzen Fessler et al. (2009) schließlich das durchschnittliche österreichische Immobilienvermögen auf 130.000 EUR (110.000 EUR ohne das Top-1% der Vermögendsten; vgl. Fessler et al. 2009: 129). Weitere Einsichten zur Vermögenssituation in Österreich werden durch Lindner (2011) mittels der OeNB Daten (2004) und Zahlen aus der EU-SILC Erhebung (2005) gegeben (vgl. Lindner 2011). Hier wird ebenfalls eine starke Ungleichverteilung von Vermögen konstatiert.

2.3. HFCS-Verteilungsstatistiken: Eine deskriptive Analyse

Im Folgenden wird eine Reihe von beispielhaften deskriptiv-statistischen Analysen präsentiert, die anhand des HFCS-Datensatzes vorgenommen werden können. Ziel ist

es, einen grundsätzlichen Eindruck über den Stand der Vermögensverteilung auf Basis der HFCS-Originaldaten zu vermitteln. In Kapitel 3 werden die hier präsentierten Graphen zum Teil erneut dargestellt, ergänzt um die davor angesprochene Korrektur der Daten am oberen Rand der Verteilungsstatistik. Die nachfolgenden Darstellungen berücksichtigen stets alle fünf Imputationen sowie die Gewichtungsfaktoren der einzelnen Beobachtungen.

Abbildung 1 zeigt die Verteilung der österreichischen Vermögen nach Vermögensklassen. Die Balken zeigen an, welcher Anteil der österreichischen Bevölkerung sich in der entsprechenden Vermögenskategorie befindet. Die oberste Vermögensklasse beginnt hier mit einem Nettovermögen (Vermögen abzüglich Schulden) von 500.000 Euro und umfasst etwas mehr als 11% der Bevölkerung.

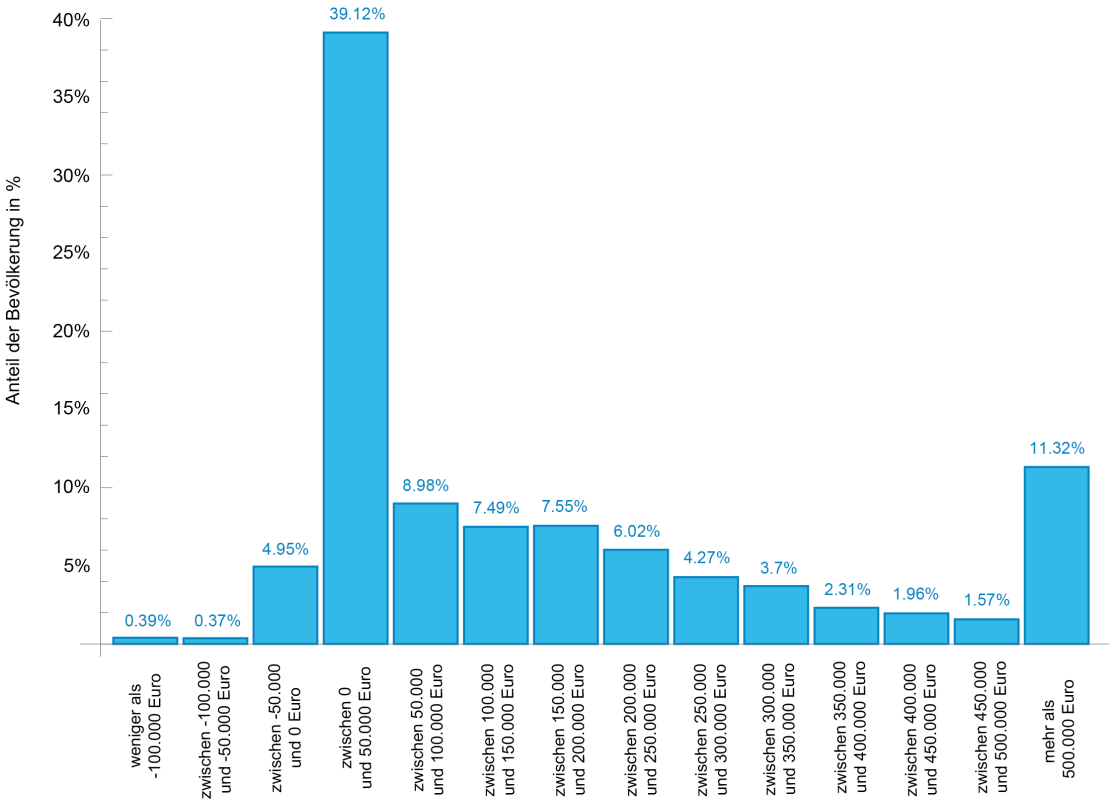


Abbildung 1: Die Vermögensverteilung in Österreich nach Vermögensklassen auf Basis der HFCS-Daten (Replikation einer OeNB-Graphik; Original in: OeNB 2012, 41)

Auf Basis der HFCS-Daten lassen sich aber freilich auch wesentlich genauere Darstellungen berechnen. Ein Beispiel hierfür ist die nachstehende Tabelle 1, die für

die Perzentile 95 bis 100 der Vermögensverteilung das kumulierte sowie das durchschnittliche Vermögen zeigt. Eine Darstellung der gesamten Perzentil-Liste findet sich in Anhang I des Forschungsberichts.

Perzentil	Gesamtvermögen im Perzentil	Durchschnittsvermögen im Perzentil
96	€ 38,815,150,674	€ 1,041,491
97	€ 48,693,272,174	€ 1,297,201
98	€ 65,450,423,948	€ 1,712,739
99	€ 94,075,818,312	€ 2,524,137
100	€ 236,958,825,570	€ 6,380,234

Tabelle 1: Vermögensverteilung der obersten 5 Perzentile

Eine andere Variante, um die Verteilung des bestehenden Vermögens abzubilden, ist es, eine kumulative Verteilungsfunktion heranzuziehen. Diese zeigt, wie sich die erwarteten Werte (d.h. die anzutreffenden Vermögen) relativ zur Bevölkerung verhalten und ermöglicht Aussagen wie *„die oberen 20% der Vermögensverteilung haben ein Nettovermögen größer als 310.000 Euro“* oder *„MillionärInnen finden sich nur in den oberen 5% der Vermögensverteilung“*. Zur leichteren Interpretation sind in der nachstehenden Abbildung einige Orientierungshilfen eingetragen.

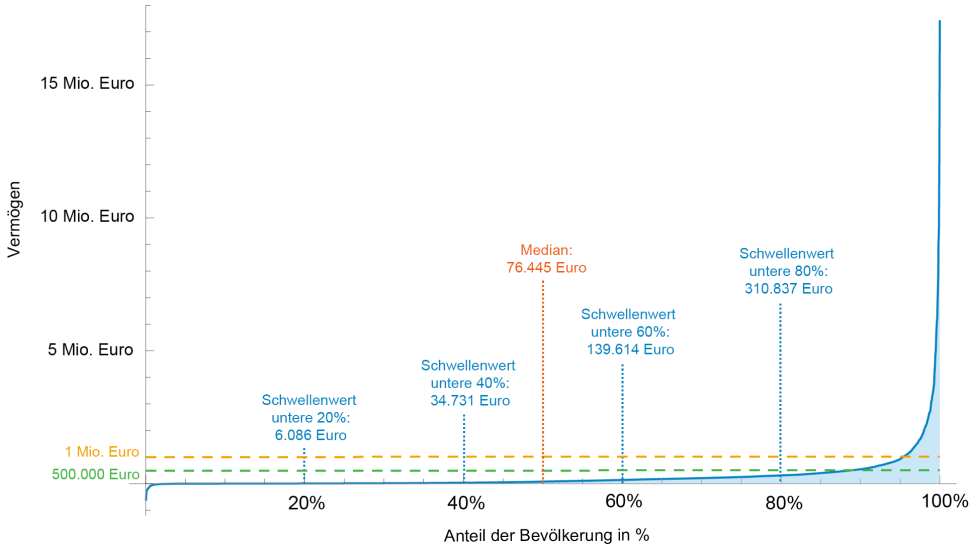


Abbildung 2: Kumulative Verteilungsfunktion der österreichischen Privatvermögen auf Basis der HFCS-Daten

Abschließend zeigt Abbildung 3 eine Lorenz-Kurve für Österreich auf Basis der HFCS-Daten. Diese Kurve wird allgemein als Indikator für wirtschaftliche Ungleichheit herangezogen und dabei auch rechnerisch in Gestalt des *Gini-Koeffizienten* ausgedrückt. Der Gini-Koeffizient ergibt sich dabei aus dem Unterschied zwischen absoluter Gleichverteilung (in der Graphik durch die 45-Grad-Linie ausgedrückt) und der realen Verteilungssituation. Die Fläche zwischen den beiden Linien repräsentiert das durch den Gini-Koeffizient gemessene Ausmaß der Ungleichheit – je größer der Wert, desto größer ist demnach auch die gemessene ökonomische Ungleichheit.

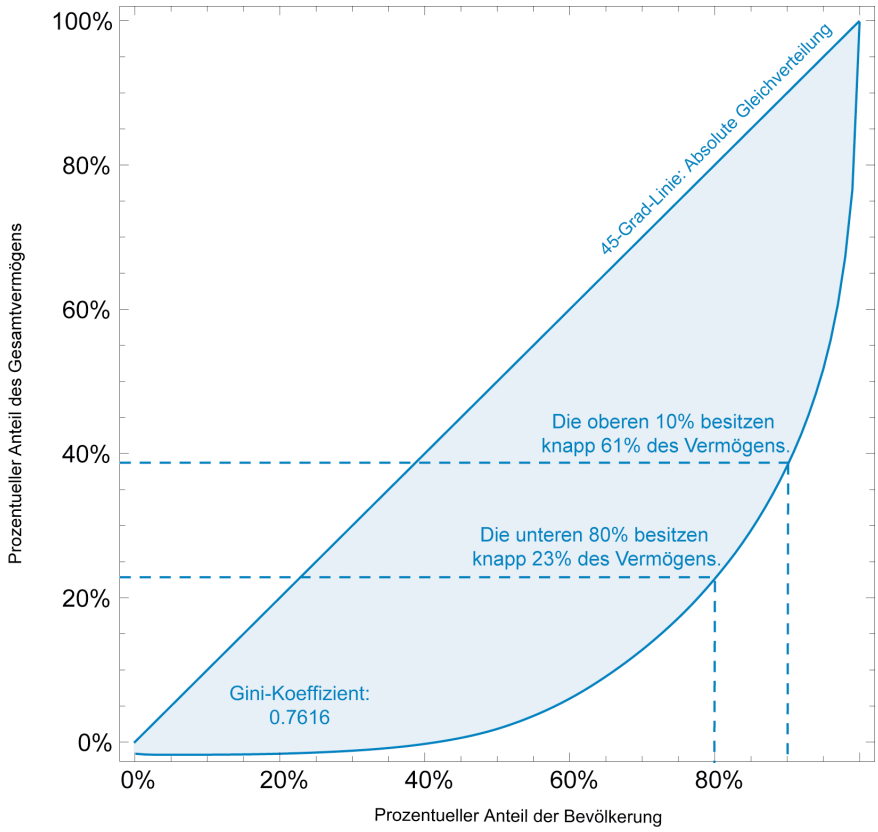


Abbildung 3: Lorenz-Kurve auf Basis der HFCS-Daten

Die nachstehende Abbildung zeigt eine zur Lorenz-Kurve vergleichbare, allerdings etwas augenfälligere Darstellung derselben Vermögensverhältnisse anhand der HFCS-Daten. Hier wird das Vermögen bestimmter Segmente der Vermögensverteilung gemäß der sich aus dem HFCS ergebenden Anteile am Gesamtvermögen abgebildet.

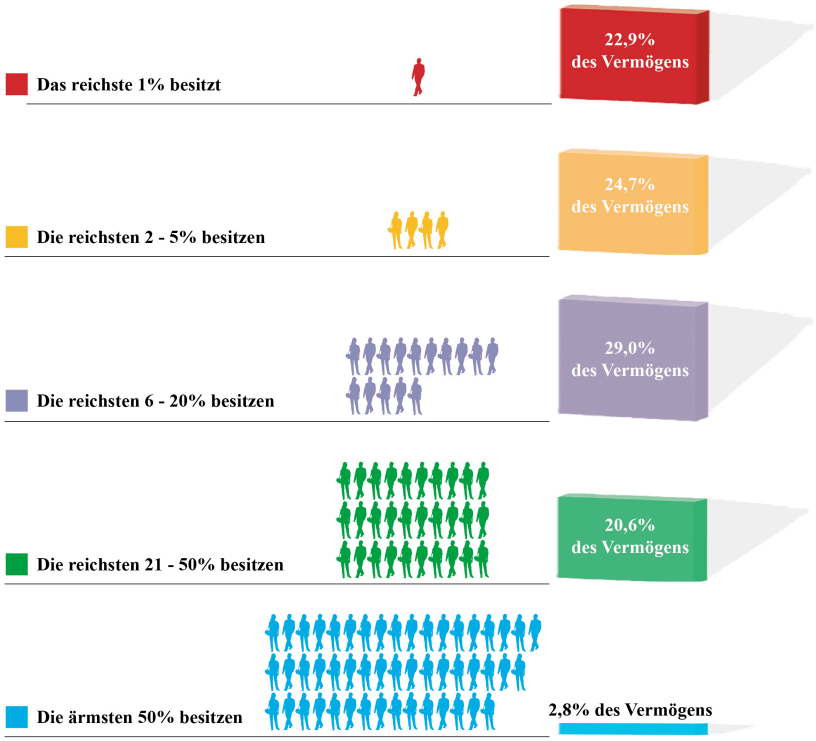


Abbildung 4: Darstellung von relativen Vermögensanteilen mittels der HFCS-Daten

3. Eine Schätzung der Verteilung des österreichischen Privatvermögens

Das spezifische Design des HFCS antizipiert eine Reihe bekannter Schwachpunkte von Vermögensbefragungen durch verschiedenste Strategien und Maßnahmen, wie etwa der Kombination aus Quotaverfahren und Zufallsziehungen, der speziellen Schulung von InterviewerInnen und der Berücksichtigung ihrer Berufserfahrung oder der Ergänzung fehlender Antworten durch statistisch aufwendige Imputationsverfahren (vgl. Albacete et al. 2012).

Weitgehend unberücksichtigt bleibt dabei allerdings der Umstand der fehlenden Abbildung der größten Vermögen im resultierenden Sample. Dabei besitzt die betreffende Personengruppe einen hohen Anteil des Gesamtvermögens und ist für eine seriöse Aufarbeitung der Vermögenssituation daher von besonderer Bedeutung.

Daher versucht die vorliegende Studie diesen Umstand ex-post zu adressieren und die in der Erhebung vorhandene Verzerrung zu kompensieren. Die angewandte Strategie

folgt dabei im Wesentlichen drei Schritten: Zuerst wird auf Basis der verfügbaren Daten für einen oberen Abschnitt der Verteilung eine Pareto-Verteilungsfunktion geschätzt. In einem zweiten Schritt werden all jene HFCS-Datensätze mit einem Nettovermögen größer als 4 Millionen Euro, also etwa ab dem obersten Vermögensperzentil, aus dem Datensatz entfernt, da in diesem Bereich zwar viele reiche Haushalte (bis rund. 20 Mio. Euro) aber eben keine besonders reichen Haushalte (mit einem Vermögen größer als 20 Mio. Euro) erfasst sind, jedoch die vorhandenen Datensätze im HFCS als für das gesamte Perzentil repräsentativ behandelt werden. In einem dritten Schritt werden Haushalte mit einem Vermögen größer als 4 Millionen Euro auf Basis der geschätzten Pareto-Verteilung generiert und dem Datensatz hinzugefügt. So wird sichergestellt, dass auch extrem reiche Haushalte im Sample vorhanden sind.

Die genaue Abfolge dieser Modifikationen des Datensatzes wird an späterer Stelle am Beispiel des HFCS-Datensatzes genau erläutert. Zuvor seien die theoretischen Grundlagen der hier vorgenommenen Transformation kurz zur Orientierung dargestellt.

3.1. Statistische Aspekte einer Theorie der Vermögensverteilung

Die Verteilungen von Einkommen und Vermögen besitzen einige Besonderheiten, die in Modellen zu ihrer Beschreibung jedenfalls Berücksichtigung finden müssen. Normalerweise zeichnen sich solche Verteilungen dadurch aus, dass sie rechtsschief sind und ein „dickes Ende“ (*fat tail*) am oberen Ende der Verteilung aufweisen, welches die Konzentration hoher Vermögens- und Einkommenswerte in den Händen weniger Individuen widerspiegelt. Korrespondierend findet sich die größte Dichte einer solchen Verteilungsfunktion im Bereich der unteren Perzentile.

3.1.1 Die Log-Normale Verteilung

Der unter anderem von Robert Gibrat (1931) vorgeschlagene Weg einer Log-Normalisierung der Daten, welcher die Anwendung vieler von einer Normalverteilung ausgehender statistischer Methoden erlauben würde, stellt den Ausgangspunkt der Überlegungen dar. Beispielsweise Cowell (2009) gibt eine ausführliche Beschreibung der Eigenschaften und Vorteile der Log-Normalen Verteilung, zu der unter anderem

die einfache Auswertung nach statistischen Standard-Methoden zählt. Klar zeigt sich jedoch, dass die Verwendung der Log-Normalen Verteilung nicht geeignet erscheint, um auch die Enden von Einkommens- und Vermögensverteilungen adäquat abzubilden (vgl. Kleiber 2007: 1; Tartal'ovà 2012).

3.1.2 Die Pareto-Verteilung

Typisch für die Verteilung von Einkommen und Vermögen ist, dass diese ab einem gewissen Schwellenwert (*Cut-off Point*) durch ein Potenzgesetz abbildbar erscheinen. Dieser ursprünglich von Vilfredo Pareto (1965[1896]) für die Einkommensverteilung Italiens identifizierte Zusammenhang wurde in zahlreichen Studien zur Schätzung des obersten Vermögenssegmentes herangezogen (vgl. u.a. Altzinger 2009; Atkinson 2006; Cowell 2011; Cowell 2009; Klass et al. 2006). Hervorzustreichen ist jedenfalls, dass Einkommen und Vermögen nur ab einem gewissen Schwellenwert paretoverteilt sind. Altzinger gibt für die Verteilung von Einkommen die obersten 20% als Richtwert an (vgl. Altzinger 2009: 7). Während die Log-Normale Verteilung eine gute Darstellung der „Mitte" der Verteilung liefert, stellt die Pareto-Verteilung eine gute Näherung für das obere Ende der Verteilung. Eine Zufallsvariable x folgt einer Pareto-Verteilung, wenn die Dichtefunktion wie folgt definiert ist, wobei α einen Formparameter und x_0 einen Skalenparameter repräsentiert (Kleiber und Klotz, 2003, 59):

$$f(x) = \frac{\alpha x_0^{\alpha}}{x^{\alpha+1}}$$

Eine manchmal in Anspruch genommene Möglichkeit zur Schätzung der Verteilungsfunktion besteht hier in der Hinzunahme von regelmäßig von verschiedensten Wirtschaftsmagazinen publizierten Reichenlisten (Forbes, Trend), um die Verteilungsparameter auf Basis dieser Top-Vermögenden zu schätzen (Atkinson 2006, Klass et al. 2006).

3.1.3 Dagum- und Singh-Madalla-Verteilungen

Der Umstand, dass sowohl die Log-Normale Verteilung, als auch die Pareto-Verteilung nur gute Näherungen für bestimmte Teile der Gesamtverteilung darstellen, führte zur Suche nach besseren Verteilungen zur Darstellung von Einkommen und

Vermögensdaten. Am bekanntesten sind dabei die von Camilo Dagum (1977) vorgeschlagene *Dagum-Verteilung*, sowie die ähnliche von Singh und Madalla vorgeschlagene *Singh-Madalla-Verteilung* (1976). Beide Verteilungen sind Spezialfälle der von McDonald (1984) vorgestellten Familie der *„generalized beta of the second kind"* (GB2) Verteilungen, die eine Dichtefunktion von

$$f(x) = \frac{ax^{ap-1}}{b^{ap}B(p,q)\left[1+(x/b)^{a}\right]^{p+q}}$$

aufweisen. Die Dagum-Verteilung als Sonderfall tritt in Erscheinung, wenn $q = 1$, die Singh-Madalla-Verteilung wenn $p = 1$. (vgl. Kleiber 1996: 266).

Während die Singh-Madalla-Verteilung, wohl aufgrund der Erstpublikation im englischsprachigen Journal „*Econometrica*" (Dagum veröffentlichte im französischsprachigen Journal „Economie Appliquèe", das wenig Verbreitung im internationalen Kontext hat), etwas bekannter ist, zeigt sich, dass die Dagum-Verteilung meist eine genauere Näherung an Vermögensdaten darstellt (vgl. Kleiber 1996; Gertel et al. 2001; Tartal'ovà 2012; Jenkins, Jäntti 2005). Eine (heuristische) Erklärung für diese Beobachtung liefert die Überlegung, dass die Dagum-Verteilung dort, wo ein Großteil der Beobachtungen von Vermögensdaten anfallen, mit zwei Formparametern ansetzt, während die Singh-Madalla-Verteilung dort nur von einem Parameter *a* determiniert wird (Kleiber 1996: 267).

3.1.4 Wahl der Verteilung im Kontext der vorliegenden Untersuchung

Für die konkret vorliegende Untersuchung wurde der Ansatz einer Pareto-Verteilung gewählt. Dies hat im Wesentlichen zwei Gründe: Zum einen eignet sich die Pareto-Verteilung aufgrund ihrer großen Bekanntheit im spezifischen Fachdiskurs ideal für eine möglichst breite Kommunikation und schnelle Verständlichkeit der erreichten Ergebnisse. Die spezifischen Vorteile der komplexeren Verteilungsfunktionen (sowohl Dagum- als auch Singh-Maddala weisen um einen Parameter mehr auf), liegen vor allem im Bereich der mittleren Vermögen, da diese konstruiert sind, um die komplette Verteilung des positiven Nettovermögens abzubilden. Insofern ist für die hier gewählte Fragestellung – die Kompensation von Datendefiziten am oberen Rand der Verteilung – eine gewisse Gleichwertigkeit der Verfahren gegeben, sofern ein

solider Ansatzpunkt für die Pareto-Verteilung vorgeschlagen werden kann. Dieser wurde im Rahmen der vorliegenden Untersuchung nicht willkürlich gewählt, sondern mit einem gezielten Verfahren bestimmt (vgl. Clauset et al. 2009). Zum anderen lieferte die Pareto-Verteilung in verschiedenen Testläufen insgesamt stabilere und daher vertrauenswürdigere Ergebnisse. Insofern wurde nach den Prinzipien Erkenntnisgewinn (etwa gleichwertig), Komplexität (Pareto-Verteilung als simplere Annahme) und Unsicherheit (Pareto-Verteilung statistisch stabiler) die Pareto-Verteilung als formaler Ausgangspunkt gewählt.

3.2. Alternative Strategien zur Quantifizierung von Top-Vermögen

Wie schon aus der oben geführten Diskussion hervorgeht, braucht es eine Strategie, um die besonders reichen Haushalte in eine adäquate Schätzung der Vermögensverteilung zu integrieren. In der Literatur wird dazu eine ganze Reihe von Wegen vorgeschlagen, die mehr oder weniger gut geeignet erscheinen, um die Informationsbasis zu erhöhen und damit die Datenlage zu verbessern. In der folgenden Analyse sollen diese auf ihre Anwendbarkeit für die Zwecke der vorliegenden Forschungsarbeit geprüft werden.

3.2.1 Reichenlisten

In zahlreichen Ländern werden meist von JournalistInnen erstellte „Reichenlisten" publiziert, welche die reichsten x Personen eines Landes namentlich anführen. Beispiele hierfür sind die Listen des Forbes-Magazins[1], in denen die national sowie international vermögendsten Personen gelistet werden, das deutsche Manager-Magazin, welches ebenfalls seit einigen Jahren eine Reichenliste erstellt[2] und das Magazin Trend, dass solche Listen für Österreich publiziert hat.[3]

Die Verwendung solcher Reichenlisten zur Verbesserung der Datenlage bzw. zur Schätzung der Verteilung von Top-Vermögen wird in zahlreichen Forschungsarbeiten angewandt (vgl. Kennickell 2000; Klass et al. 2006), obwohl die Aussagekraft dieser Listen bestenfalls beschränkt ist. Atkinson (2006) notiert dazu, dass die Erstellung einer solchen Liste durch JournalistInnen auch bei guter Recherche nur eine Näherung

[1] vgl.: http://www.forbes.com/billionaires/ bzw. http://www.forbes.com/forbes-400/list/
[2] vgl.: http://www.manager-magazin.de/unternehmen/artikel/a-860164.html
[3] vgl.: http://www.trendtop500.at/die-reichsten-oesterreicher/

darstellen kann: JournalistInnen sind bei der Erstellung solcher Listen ebenso von Auskünften über die Vermögenslage des obersten Segments angewiesen wie Befragungen und damit mit ähnlichen Problemen der Antwortverweigerung konfrontiert. Weiters ergeben sich auch bei Vorliegen von Detailinformationen zahlreiche Bewertungsschwierigkeiten. Schließlich ist die Schätzung von Schulden, die vom Bruttovermögen abzuziehen sind, um eine Einschätzung zu erhalten, meist noch schwieriger als die Einschätzung der Vermögensbestandteile. Nicht zuletzt existiert noch das Problem, dass zahlreiche Angaben in den Reichenlisten (siehe die Trend-Liste) ganze Familien-Clans als TrägerInnen von Unternehmensbesitz oder Erbschaften anführen, was die Analyse solcher Reichenlisten beträchtlich erschwert (vgl. Atkinson 2006:6).

Die Schätzung der Vermögensverteilung im obersten Segment mittels einer Reichenliste benötigt das Vorliegen einer Pareto-Verteilung, die folgenden Zusammenhang zwischen Rangzahl einer Person (r; r=1 für die reichste Person) und ihrem Vermögen (w) herstellt:

$$r = e^{A} w_r^{-\alpha}$$

Wobei w_r das Vermögen einer Person mit Rangzahl r bezeichnet, A eine Konstante und α der Pareto-Exponent ist. Dieser Zusammenhang hat eine einfache graphische Repräsentation: Logarithmiert man obigen Ausdruck ($\log r = A - \alpha \log w_r$) und trägt das (logarithmierte) Nettovermögen (w) auf der horizontalen Achse und die (logarithmierte) Rangzahl auf der vertikalen Achse auf, so ergibt sich eine Gerade mit der Steigung α. Sollte also tatsächlich eine Pareto-Verteilung vorliegen, so sollten sich diese Personen etwa auf einer Linie befinden (vgl. Kleiber/Klotz, 2003: 61).

Mittels einer OLS-Schätzung lässt sich daraus unter Verwendung der Trend Reichenliste ein Pareto-Alpha von 1,3074 errechnen, was mit den Schätzungen in anderen Ländern durchaus konsistent ist. Darüber hinaus ist diese Schätzung ebenfalls konsistent mit den aus dem HFCS-Datensatz errechneten Alphas, welche über alle Implicates einen Mittelwert von 1,28 aufweisen. Nichtsdestotrotz ist die Methode der Pareto-Alpha-Schätzung auf Basis einer Reichenliste mit zu vielen Unsicherheiten behaftet und kann nur als erste Näherung betrachtet werden. Die relativ gute

Übereinstimmung der Werte spricht aber tendenziell dafür, dass die hier bestimmte Verteilung auch die obersten Vermögen zuverlässig beschreibt.

3.2.2 Weitere Datenquellen

Zusätzliche in der Literatur vorgeschlagene Datenquellen sind in direkte Quellen (wie etwa Vermögenssteuerdaten) und indirekte Quellen (wie Daten zu Immobilien oder Einkommen) zu unterscheiden. Da eine direkte Vermögensbesteuerung in Österreich nicht vorgesehen ist, fällt die erstgenannte Variante weg. Nichtdestotrotz wäre aber auch eine solche Vorgehensweise mit zahlreichen Risiken verbunden, da die Erfassung in Statistiken zur Vermögensbesteuerung natürlich von der Konzeption der Steuer abhängig ist. So finden sich am Beispiel von Frankreich, wo eine durchgängige Vermögenssteuer eingehoben wird, zahlreiche Möglichkeiten zur Herabsetzung der Bemessungsgrundlage und Ausnahmeregelung in der Erfassung sowie Bewertung bestimmter Vermögensbestandteile, die eine einfache Übernahme der Vermögenssteuerdaten nicht möglich machen. Atkinson (2006) merkt darüber hinaus an, dass es auch durch die generelle Tendenz der Steuervermeidung zu beträchtlichen Verzerrungen kommen kann (vgl. Atkinson 2006: 7).

Eine Berechnung der Vermögensverteilung für das oberste Segment auf Basis von Erbschafts- und Schenkungssteuern fällt für Österreich ebenfalls insofern weg, als dass solche Steuern seit 1. August 2008 nicht mehr eingehoben werden. Die Vorgehensweise bei einer solchen Berechnung geht jedoch ohnehin ebenso mit zahlreichen statistischen Unwegsamkeiten in der Bewertung der Vermögen einher. Vor allem für Österreich wäre nach dem alten Modell der Erbschafts- und Schenkungssteuer von einer massiven Unterschätzung der wahren Vermögenssituation auszugehen, da den Daten der Ansatz der Einheitsbewertung zugrunde gelegt wäre, welche oftmals den wahren Wert der Grundstücke nicht mehr wiedergeben.

Abschließend lässt sich festhalten, dass alle drei oben vorgestellten Methoden zur Schätzung der Vermögensverteilung unter den besonders reichen Haushalten mit methodischen und praktischen Unsicherheiten behaftet sind, welche die Genauigkeit der Schätzung vermindern. Darüber hinaus zeigt sich vor allem für den österreichischen Fall, dass die Datenlage als unzureichend zu bewerten ist, um die

oben vorgestellten Verfahren einzusetzen.

Die Details zur von uns angewandten Methode der Parameterschätzung der Pareto-Verteilung werden in den folgenden Abschnitten genauer dargestellt. Der dazugehörige Rechengang kann zusätzlich mit dem in Anhang II befindlichen Mathematica-Code repliziert werden. Die sich daraus ergebenden Änderungen werden in Kapitel 3.4 und Anhang III (Perzentilliste auf Basis der modifizierten Daten) dargestellt.

3.3. Schätzverfahren und Datenbearbeitung

Die Schritte der nachfolgenden Datenkorrektur lassen sich nun folgendermaßen zusammenfassen: Zuerst wird mithilfe des Cramer-von-Mises Tests ein geeigneter Ansatzpunkt für die Pareto-Verteilung bestimmt. Die Daten oberhalb dieses Ansatzpunktes bilden dann die Datenbasis für die Schätzung der Pareto-Verteilung. Schließlich werden alle Haushalte mit einem Vermögen > 4 Millionen (also der zu korrigierende Teil) aus dem Datensatz entfernt und durch neu generierte Haushalte ersetzt, welche aus einer Serie von Zufallsziehungen aus dem oberen Bereich (> 4 Millionen) der zuvor geschätzten Pareto-Verteilung gewonnen werden. Die nachstehende Abbildung 5 fasst die hier implementierte Vorgangsweise schematisch zusammen, während die folgenden Teilkapitel die technische Vorgehensweise im Detail beschreiben.

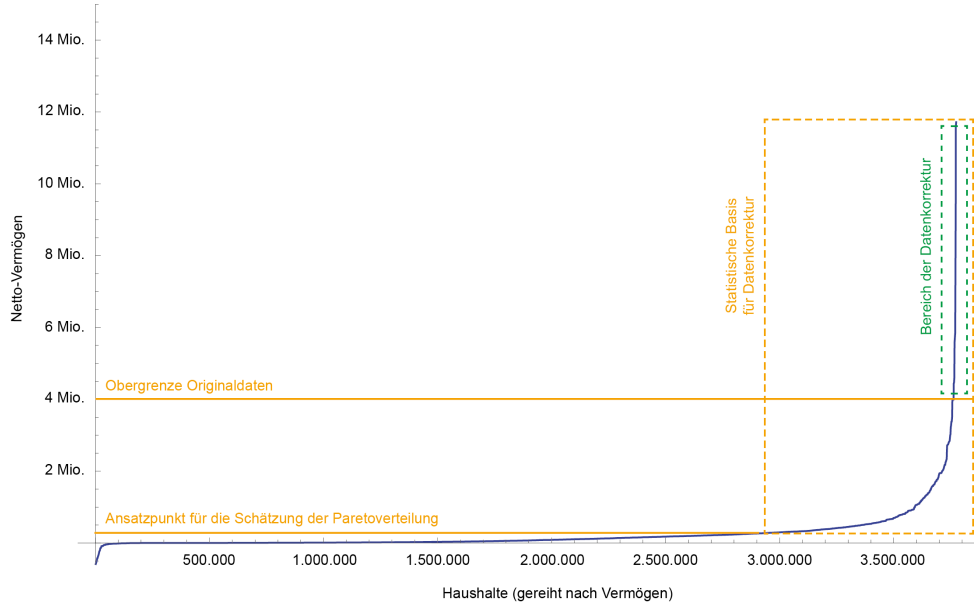


Abbildung 5: Veranschaulichung der Methode zur Daten-Korrektur

3.3.1 Bestimmung der Verteilungsparameter

Bei der Anwendung der Pareto-Verteilung ist die Wahl eines geeigneten Ansatzpunktes für die Verteilung von zentraler Relevanz. Gesucht wird also jener Punkt x für den gilt, dass die Vermögenswerte > x paretoverteilt sind. Obgleich die Wahl dieses Ansatzpunktes eine beträchtliche Implikation für die Ergebnisse mit sich bringen kann, wird dieser nur in seltenen Fällen mit Hilfe statistischer Methoden bestimmt, sondern zumeist der Literatur oder theoretischen Erwägungen entlehnt (vgl. beispielhaft Bach et al. 2010, Bach/Beznoska 2012).

In der hier vorliegenden Studie wurde primär die erste Strategie verfolgt und der relativ geeignetste Ansatzpunkt isoliert. Ein Mathematica-Programm zur Durchführung einer solchen Berechnung findet sich in Anhang II. Im Detail wurde diese Operation wie folgt durchgeführt:

In einem ersten Schritt wurden die Schwellenwerte für die Perzentile 71-100 als mögliche Ansatzpunkte gewählt und über alle fünf Imputationen mit Hilfe eines Maximum Likelihood Schätzers (Clauset et al. 2009) die entsprechenden Parameterwerte errechnet. Der jeweilige Ansatzpunkt (x_0) ergab sich hier also aus dem Schwellenwert für das jeweilige Perzentil und die entsprechenden Alpha-Parameter zur Komplettierung der Pareto-Verteilung wurden aus der sich jeweils ergebenden Grundgesamtheit geschätzt. Die resultierenden Verteilungen wurden in einem zweiten Schritt mittels einem für diese Fragestellung geeignetem statistischen Test, dem Cramer-von-Mises-Test, auf ihre Plausibilität in Relation zu den zugrundeliegenden Daten hin überprüft. Eine wichtige methodische Anmerkung in diesem Kontext ist, dass die Nullhypothese dieser Tests annimmt, die Daten entsprächen einer Pareto-Verteilung mit den jeweiligen Parametern. Gemäß dieser Interpretation sind hohe p-Werte bei den Tests als Signum für die entsprechende Verteilung auszulegen, da in diesen Fällen die Nullhypothese nicht verworfen werden kann.

Die nachstehende Abbildung zeigt einen entsprechenden Testdurchlauf über alle fünf Imputationen mittels des Cramer-von-Mises-Tests. Der obere Teil der Abbildung zeigt den Verlauf der Schätzwerte für Paretos-Alpha entlang der letzten dreißig Perzentile, der untere Teil der Abbildung liefert die korrespondierenden p-Werte.

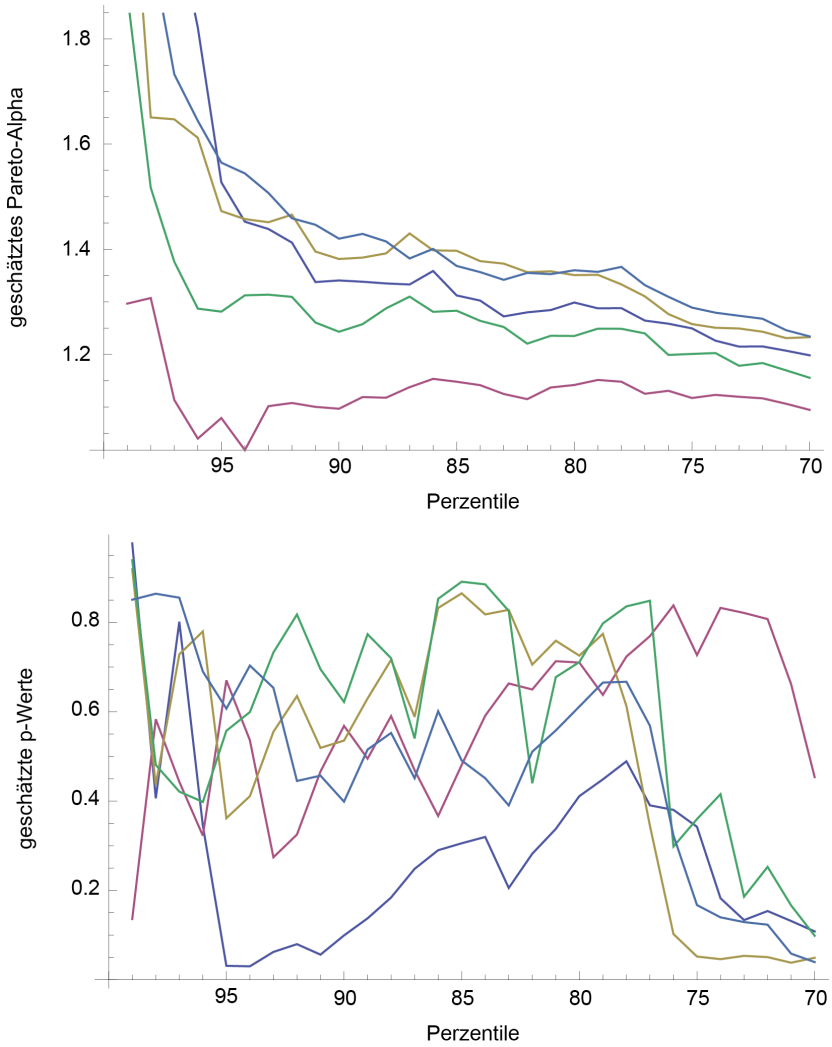


Abbildung 6: Geschätzte Alpha-Parameter und dazugehörige p-Werte (nach Cramer-von-Mises) für die obersten 30 Perzentilgrenzen (unter Berücksichtigung aller Imputationen)

Hier wird ersichtlich, dass gerade die Daten am oberen Rand der Verteilung als Ansatzpunkte für eine formale Modellierung nur wenig vertrauenswürdig scheinen, da hier die statistischen Resultate stark mit der jeweils verwendeten Grundgesamtheit schwanken. Umgekehrt gibt es Bereiche stabilerer Schätzungen (etwa rund um das 80. Perzentil), in denen sich die Schwankungen über die Grundgesamtheit bzw. zwischen den Imputationen innerhalb eines relativ stabilen Korridors einpendeln. Die nachstehende Tabelle zeigt die geschätzten Werte für die Pareto-Alphas sowie die Nettovermögen der jeweiligen Ansatzpunkte (Schwellenwert des 78. Perzentils) über alle fünf Imputationen.

Imputation	Paretos Alpha	Ansatzpunkt für Pareto-Verteilung
#1	1.28808	Nettovermögen von 281.242 Euro
#2	1.14815	Nettovermögen von 287.809 Euro
#3	1.3332	Nettovermögen von 289.811 Euro
#4	1.24881	Nettovermögen von 293.161 Euro
#5	1.36649	Nettovermögen von 288.422 Euro
Durchschnitt	1.276946	Nettovermögen von 288.089 Euro

Tabelle 2: Nettovermögen und geschätzte Pareto-Alphas am Schwellenwert des 78. Perzentils

Der konkrete Ansatzpunkt für die Pareto-Verteilung wurde schließlich in zwei Schritten gewählt. Zuerst wurde ein Intervall bestimmt, in dem der Test über alle fünf Imputationen p-Werte größer als 0,2 erreichte. Um in diesem Intervall den aus statistischer Sicht vertrauenswürdigsten Ansatzpunkt zu bestimmen, wurde in Folge nach jenem Punkt gesucht, in dem das Minimum der fünf p-Werte im jeweiligen Perzentil maximal ist. Für den Cramer-von-Mises-Test ist dies das 78. Perzentil. Dieses wurde in der Folge als Ansatzpunkt mit der geringsten statistischen Unsicherheit identifiziert und als Basis für alle weiteren Berechnungen verwendet.

3.3.2 Elimination und Ergänzung von Haushalten

Die Korrektur des vorhandenen Datensatzes erfolgte in insgesamt vier Schritten, wobei sich in Anhang IV der entsprechende verwendete Mathematica-Code befindet. Zuerst werden dabei all jene Beobachtungen im HFCS eliminiert, die ein Nettovermögen von über 4 Millionen Euro aufweisen. Durch diese Maßnahme gehen relativ wenig Haushalte verloren (je nach Imputation zwischen 8 und 30 Beobachtungen, die zwischen 11.374 und 44.081 Haushalte repräsentieren), wobei man hier analog zum bereits Gesagten annimmt, dass in diesen Beobachtungen besonders reiche Haushalte nicht bzw. nicht ausreichend abgebildet sind.

In einem zweiten Schritt wird die jeweilige kumulative Pareto-Verteilungsfunktion mit den zuvor geschätzten Parametern herangezogen und auf Basis einer simplen Schlussrechnung die Zahl der Haushalte errechnet, die ein Vermögen größer als 4

Millionen Euro aufweisen. Hierzu wird aus dem HFCS-Datensatz die Anzahl der Haushalte zwischen dem jeweiligen Ansatzpunkt für die Pareto-Verteilung und der 4-Millionen-Grenze entnommen und in die entsprechende Formel eingefügt. Die entsprechende kumulative Dichtefunktion (hier: CDF) liefert den Prozentanteil all jener Haushalte, die sich *über* dem Ansatzpunkt, aber *unterhalb* der 4-Millionen-Grenze befinden. Daraus ergibt sich die folgende Schlussrechnung zur Bestimmung der Zahl der Haushalte mit einem Nettovermögen größer als 4 Millionen Euro (X = Zahl der zu ergänzenden Haushalte, Y = Zahl der Haushalte zwischen Ansatzpunkt und 4 Millionen).

$$X = \frac{Y \cdot (1 - CDF)}{CDF}$$

Diese Vorgangsweise impliziert, dass zur Ergänzung fehlender Haushalte nur auf die qualitativ hochwertigen Daten aus der HFCS-Erhebung zurückgegriffen werden muss und – im Gegensatz zu anderen derartigen Ansätzen (vgl. Bach et al. 2010, Bach/ Beznoska 2012) – nicht relativ viel weniger vertrauenswürdige Quellen (z.B. Reichenlisten) zur Durchführung der entsprechenden Berechnungen verwendet werden.

Der nächste Schritt erfordert die berechnete Anzahl von Haushalten, die sich am oberen Ende der Vermögensskala befinden, aber von den Daten nicht mehr repräsentiert werden (also jene Haushalte mit einem Nettovermögen größer als 4 Millionen Euro), aus der Verteilungsfunktion zu generieren. Zu diesem Zweck wird der gewichtete Durchschnitt aus fünf Ziehungen aus der geschätzten Pareto-Verteilung berechnet, wobei die Anzahl der zu ziehenden Haushalte von der Zahl neu zu generierender Haushalte bestimmt wird, welche wiederum von Imputation zu Imputation unterschiedlich sind. Diese werden in Folge dem HFCS-Datensatz mit einem Gewicht von 1 hinzugefügt. Jede der (gewichteten) Ziehung repräsentiert damit exakt einen Haushalt.

Zuletzt müssen die - durch die vorgenommene Modifikation nicht mehr 100% kohärenten - Gewichtungen der Original-Haushalte korrigiert werden. Dabei wird die Nettoveränderung innerhalb der jeweiligen Imputation (diese bewegt sich zwischen +17.000 und – 4.000 Haushalten) in Relation zur jeweiligen Gesamtbevölkerung

gesetzt und zur Korrektur die jeweiligen Gewichte linear abgeschmolzen bzw. aufgewertet.[4]

Einen letzten nennenswerten Faktor bildet die vorgenommene „Deckelung" der Zufallsziehung. Die zufällig zu ziehenden Vermögenswerte wurden mit *einer Milliarde Euro* gedeckelt. Obgleich das Modell eine durchaus realistische Anzahl an MilliardärInnen produziert, haben wir uns entschieden, die entsprechende Ziehung zu begrenzen, um allzu drastische Ausreißer nach oben gänzlich zu vermeiden. Anders formuliert: Die verwendeten Samples enthalten aus Gründen der kalkulatorischen Vorsicht keine Milliardäre oder Milliardärinnen.

3.4. Die Vermögensverteilung in Österreich unter Berücksichtigung der Datenkorrektur

Eine erste zentrale Frage hinsichtlich der hier vorgenommenen Datenmodifikation ist die Auswirkung der Berücksichtigung der Top-Vermögen auf die Gesamtstruktur des Samples. Um diesen Aspekt zu analysieren, zeigt Abbildung 7 - in Analogie zu Abbildung 1 - wiederum die auf Vermögensklassen aufgeteilte Gesamtbevölkerung. Veränderungen in den Werten bzw. der Darstellung werden im Folgenden farblich hervorgehoben. In dieser Variante der aggregierten Darstellung ist nur eine minimale Veränderung erkennbar. Zu sehen ist lediglich, dass sich die entsprechenden Bevölkerungsanteile ganz leicht nach unten verschieben, bis auf die Klasse mit über €500.000 Nettovermögen. Deren Anteil steigt aufgrund der hinzugefügten besonders reichen Haushalte leicht an.

[4] Für die Implicates 1, 3, 4 und 5 ist die Anzahl der neu hinzugefügten Haushalte größer als jene der eliminierten Haushalte (jene mit einem Vermögen über 4 Millionen Euro). Hier müssen also die Gewichte der im Datensatz verbleibenden Haushalte proportional abgeschmolzen werden. In Implicate 2 ist die Anzahl der eliminierten Haushalte größer als die Zahl der neu generierten Haushalte. Dies ergibt sich aus dem Umstand, dass im Vergleich zu den anderen Implicates die Haushalte mit einem Vermögen über 4 Millionen deutlich höher ist und dadurch eine größere Zahl an Haushalten aus dem Sample entfernt werden muss. In der Folge müssen hier also die Gewichte der verbleibenden Haushalte proportional aufgewertet werden.

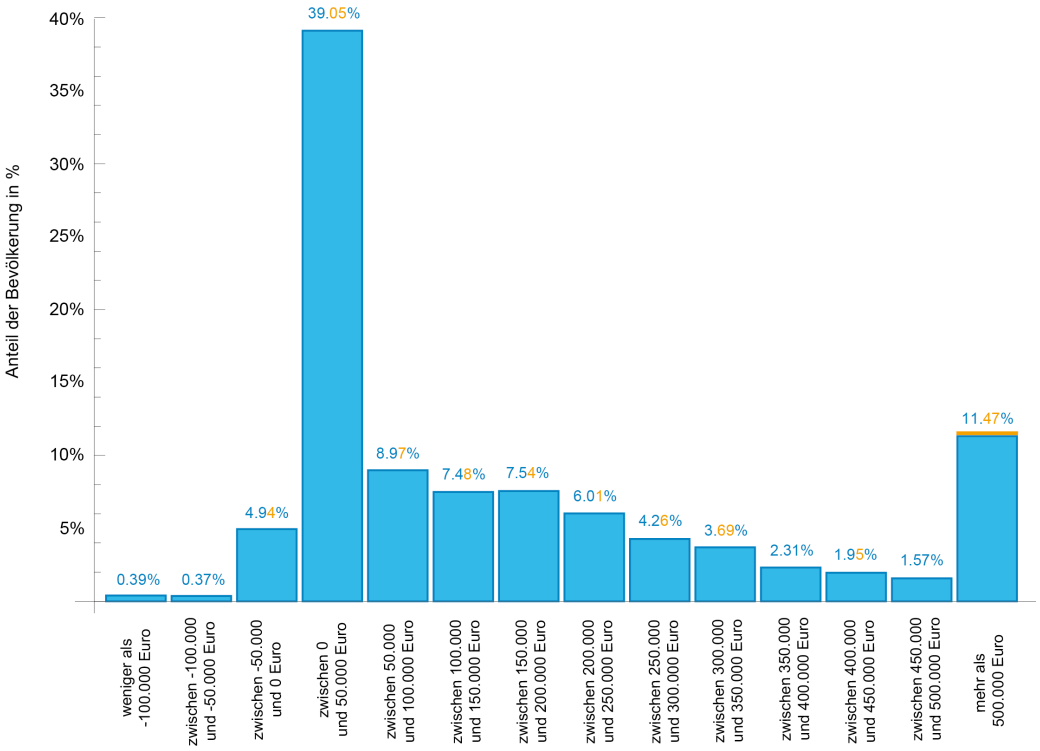


Abbildung 7: Die Vermögensverteilung in Österreich nach Vermögensklassen auf Basis der *modifizierten* HFCS-Daten (Replikation von Abbildung 1; Unterschiede in Orange)

Auch eine Replikation der Perzentilliste wie in Tabelle 1 bringt vergleichbare Ergebnisse: besonders eklatant ist nur der Vermögensstand des reichsten Perzentils, nämlich um 98,6% gewachsen. Damit steigt auch das sich aus der Schätzung ergebende gesamte Nettovermögen signifikant von etwa 1000 Mrd. Euro auf etwa 1249 Mrd. Euro. Die nachstehende Tabelle 3 zeigt die Daten der obersten 5 Perzentile. Die gesamte modifizierte Perzentilliste findet sich in Anhang III des Forschungsberichtes[5].

[5] Ein potentieller Weg um die Plausibilität unserer Berechnungen zu prüfen, besteht in einem Vergleich des Finanzvermögens laut HFCS mit dem Finanzvermögen laut Gesamtwirtschaftlicher Finanzierungsrechnung (*GFR*). Ein solcher Vergleich ist allerdings aufgrund einiger definitorischer Unterschiede aber kaum möglich. So werden in der GFR in der Rubrik *Aktien* alle Unternehmensbeteiligungen erfasst, während im HFCS nur jene Unternehmensbeteiligungen als Aktien gezählt werden, bei denen kein Haushaltsmitglied in der Unternehmensführung mitwirkt. Liegt ein solches Mitwirken vor, so fällt es im HFCS in die Rubrik *Firmenvermögen* und wird nicht mehr zum Finanz-, sondern zum Sachvermögen gezählt. Ein weiteres Beispiel ist die unterschiedliche Erfassung von Pensionsvermögen. Siehe hierzu auch OeNB (2013).

Perzentil	Gesamtvermögen im Perzentil	durchschnittliches Vermögen im Perzentil
96	€ 40,397,367,382	€ 1,073,117
97	€ 50,791,826,986	€ 1,341,788
98	€ 66,601,914,082	€ 1,768,638
99	€ 101,014,993,783	€ 2,690,588
100	€ 469,058,597,640	€ 12,670,045

Tabelle 3: Die Vermögen der reichsten fünf Perzentile auf Basis der modifizierten HFCS-Daten.

Analog zu Kapitel 2 zeigt Abbildung 8 die modifizierte Lorenz-Kurve anhand derer die Erfassung der besonders reichen Haushalte und die Auswirkungen auf die Vermögensverteilung gut sichtbar dargestellt werden kann.

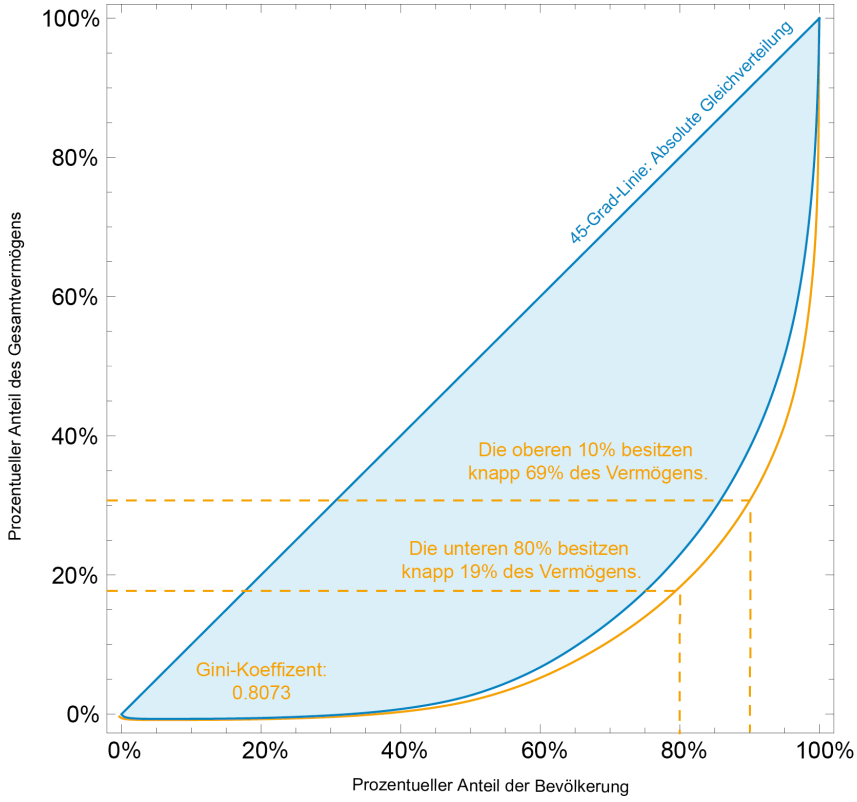


Abbildung 8: Lorenz-Kurve mit modifizierten HFCS-Daten

Wiederum analog zu Abbildung 4 zeigt die nächste Abbildung die veränderten Vermögensverhältnisse anhand des Anteils verschiedener Vermögensgruppen. Beachtenswert ist hier, dass der Anteil des obersten Prozents von 22,9% auf 37% des Gesamtvermögens ansteigt.

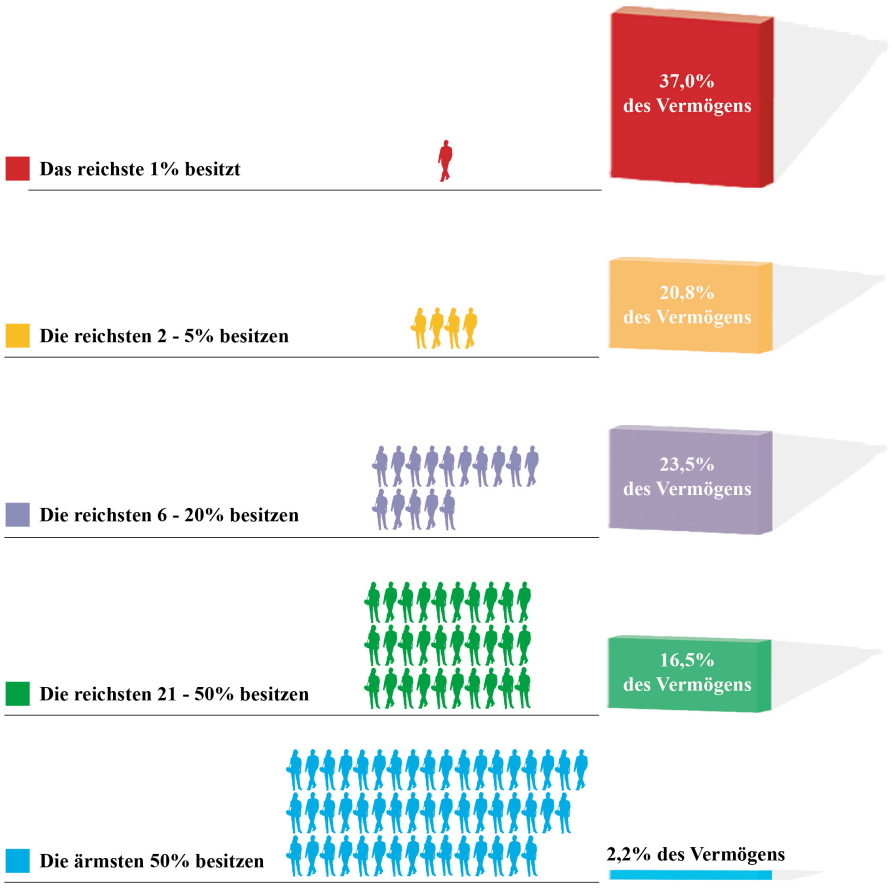


Abbildung 9: Darstellung von relativen Vermögensanteilen mittels der korrigierten HFCS-Daten

4. Resümee

In der vorliegenden Arbeit wurde versucht, die fehlende Erfassung besonders reicher Haushalte im HFCS und die sich dadurch ergebende Verzerrung zu korrigieren. Von den erprobten Verfahren (Pareto-, Dagum und Singh-Madalla-Verteilung sowie Zuhilfenahme der Trend-Liste der reichsten ÖsterreicherInnen) stellte sich die Schätzung des obersten Bereichs der Vermögensverteilung mittels einer Pareto-Verteilung als die einfachste und zugleich stabilste Variante heraus. Für die

Schätzung der Pareto-Verteilung wurde dabei etwas mehr als das obere Fünftel der HFCS-Daten herangezogen. Der genaue Ansatzpunkt der Verteilung wurde dabei mit Hilfe eines geeigneten statistischen Tests (Cramer-von-Mises) bestimmt. Anschließend wird der vorhandene HFCS-Datensatz durch Zufallsziehungen aus dieser Verteilung ergänzt, um das oberste Vermögenssegment (Nettovermögen > 4 Millionen Euro) besser abbilden zu können. Zugleich werden alle im HFCS vorhandenen Haushalte mit einem Vermögen über 4 Millionen entfernt.

Es zeigt sich, dass eine solche Korrektur erhebliche Auswirkungen auf die Ergebnisse hat. So steigt dadurch das geschätzte Gesamtvermögen von etwa 1000 Mrd. EUR auf 1249 Mrd. EUR, wovon der Großteil des Zugewinns auf das oberste Perzentil entfällt (das Vermögen dieses Perzentils steigt um 98,6%). Daraus ergibt sich unter anderem, dass die reichsten 10% der ÖsterreicherInnen nicht 61% (HFCS) sondern 69% des Gesamtvermögens besitzen.

5. Literatur

Albacete, N./ Lindner, P./ Wagner, K./ Zottel, S. (2012): Household Finance and Consumption Survey des Eurosystems 2010: Methodische Grundlagen für Österreich. *Geldpolitik & Wirtschaft*, 3/2012-Addendum.

Altzinger, W. (2009): Die Entwicklung der Spitzeneinkommen in Österreich. Online: http://www.wiwiss.fu-berlin.de/forschung/veranstaltungen/rse/papers_winter_09_10/paper_altzinger.pdf

Andreasch, M./ Fessler, P./ Schürz, M. (2009): Unternehmensbeteiligungen der privaten Haushalte in Österreich. *Geldpolitik & Wirtschaft*, 4/2009.: 66-84.

Atkinson, A.B. (2006): Concentration Among the Rich, Research Paper No. 2006/151, World Institute for Developement Economics Research.

Atkinson, A./ Piketty, T./ Saez, E. (2011): Top incomes in the long run of history. *Journal of Economic Literature* 49(1): 3–71.

Avery, R.B./ Elliehausen, G.E./ Kennickell, A.B. (1986): Measuring wealth with survey data: an evaluation oft the 1983 survey of consumer finances, *Review of Income and Wealth* 34(4): 339-369.

Bach, S./Beznoska, M./Steiner, V. (2012): Aufkommens- und Verteilungswirkung einer Grünen Vermögensabgab. DIW: Politikberatung Kompakt 59, Berlin.

Bach, S./Beznoska, M. (2012): Aufkommens- und Verteilungswirkung einer Wiederbelebung der Vermögenssteuer. DIW: Politikberatung Kompakt 68, Berlin.

Beer, C./ Mooslechner, P./ Schürz, M./ Wagner, K. (2006): Das Geldvermögen privater Haushalte in Österreich: eine Analyse auf Basis von Mikrodaten, *Geldpolitik und Wirtschaft*, 2/2006: 101-119.

Borgerhoff Mulder, M./ Bowles, S./ Hertz, T./ Bell, A./ Beise, J., et al. (2009): Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. Science 326(5953): 682-688.

Clauset, C./ Shalizi, C. R./ Newman, M. E. J. (2009): Power Law Distributions in Empirical Data. SIAM Review 51(4): 661-703.

Clementi, F/ Gallegati, M./ Kaniadakis, G. (2012): A generalized statistical model for the size distribution of wealth. *Journal of Statistical Mechanics: Theory and Experiment* 2012(12), P12006.

Cowell, F.A. (2009): *Measuring Inequality.* Oxford University Press.

Cowell, F.A. (2011): Inequality Among the Wealthy. Centre for Analysis of Social Exclusion at the London School of Economics.

Dagum, C. (1977): A new model of personal income distribution: Specification and estimation, *Economie Appliquée* 30: 413-437.

Fessler, P./ Mooslechner, P./ Schürz, M./ Wagner, K. (2009): Das Immobilienvermögen privater Haushalte in Österreich. *Geldpolitik & Wirtschaft*,

2/2009: 113-134.

Fessler, P/ Mooslechner, P./ Schürz, M. (2012): Household Finance and Consumption Survey des Eurosystems 2010: Erste Ergebniss für Österreich. *Geldpolitik und Wirtschaft*, 3/2012: 26-67.

Frick, J. R./ Grabka, M. (2009): Gestiegene Vermögensungleichheit in Deutschland. Wochenbericht des DIW Berlin Nr. 4/2009.

Gertel, H.R/ Giuliodori, R./ Auerbach, P. F/ Rodrìguez, A. F (2001): An estimation of personal income distribution using the Dagum Model with an application to Cordoba 1992-2000. Online: http://cdi.mecon.gov.ar/biblio/docelec/utdt/pob2-5.pdf.

Gibrat, R. (1931): *Les Inégalités économiques*. Librairie du Recueil Sirey Paris

Guger, A./ Marterbauer, M. (2007): Langfristige Tendenzen der Einkommensverteilung in Österreich – ein Update. WIFO Working Papers, Nr. 307.

Guttmann, R./ Philon, D. (2010): Consumer debt and financial fragility. *International Review of Applied Economics* 24(3): 269-283

Hahn, F.R./ Magerl, C. (2006): Vermögen in Österreich; WIFO Monatsbericht 1/2006.

Hoeller, P./ Joumard, I./ Bloch, D./ Pisu, M. (2012): Less Income Inequality and More Growth - Are They Compatible?: Part1: Mapping Income Inequality Across the OECD. OECD Working Paper No. 924.

Jenkins, S.P/ Jäntti, M. (2005): Methods for Summarizing and Comparing Wealth Distributions. ISER Working Paper Number 2005-05.

Kennickell, A.B. (2000): An Examination of Changes in the Distribution of Wealth From 1989 to 1998: Evidence from Consumer Finances. Working Paper No. 307 (Federal Reserve Board).

Klass, O.S./ Biham, O./ Levy, M./ Malcai, O./ Solomon, S. (2006): The Forbes 400 and the Pareto wealth distribution. *Economics Letters* 90(2): 290-295.

Kleiber, C./ Kotz, S. (2003): *Statistical Size Distribution in Economics and Actuarial Sciences*. Hoboken: Wiley-Interscience.

Kleiber, C. (1996): Dagum vs. Singh-Maddala income distributions. *Economics Letters* 53(3): 265-268.

Kleiber, C. (2007): A Guide to the Dagum Distributions. WWZ Working Paper 23/07

Lindner, P. (2011): Decomposition of Wealth and Income using Micro Data from Austria. Österreichische Nationalbank; Working Paper 173.

Little, R.J.A./ Rubin, D.B. 2002. Statistical Analysis with Missing Data. New York: Wiley. Zweite Auflage.

Piketty, T./ Saez, E. (2006): The Evolution of Top Incomes: A Historical and International Perspective. *American Economic Review* 96(2): 200-205.

McDonald, J.B. (1984): Some generalized functions for the size distribution of

income. *Econometrica*, 52(3): 647-663.

OeNB (2013): Sektorale VGR in Österreich 2012: Integrierte Darstellung der Wirtschafts- und Finanzkonten für Haushalte, Unternehmen, Staat und Finanzsektor in den Volkswirtschaftlichen Gesamtrechnungen; Statistiken (Sonderheft), Juni 2013.

Pareto, V. (1965[1896]): La Courbe de la Repartiation de la Richesse. In G. Busino (ed.): *Oevres Completes de Vilfredo Pareto*, Genf: Librairie Droz.

Schimpel, M. (2004): Reimplementierungsmöglichkeiten der persönlichen allgemeinen Vermögensteuer in Österreich. Diplomarbeit am Institut für Finanzwissenschaft der Wirtschaftsuniversität Wien.

Singh, S.K./ Madalla, G.S. (1976): A function for the size distribution of income. *Econometrica*,. 44: 963-970.

Stiglitz, J. E. (2012): *The price of inequality: How today's divided society endangers our future*. New York, London, WW Norton & Company

Tartal'ovà, A. (2012): Modelling Income Distribution in Slovakia. The 6th International Days of Statistics and Economics, Prague, September 13-15.

Wilkinson, R. G. und Picket K. E. (2007): The problems of relative deprivation: why some societies do better than others. *Social Science & Medicine* 65(9): 1965-1978.

Anhang I: Perzentilliste auf Basis der HFCS-Daten

Perz-entil	Gesamtvermögen im Perzentil	Durchschnittsvermögen im Perzentil	Perz-entil	Gesamtvermögen im Perzentil	Durchschnittsvermögen im Perzentil
1	-€ 5,500,714,505	-€ 143,347	51	€ 3,059,497,442	€ 80,111
2	-€ 852,806,627	-€ 22,396	52	€ 3,291,553,550	€ 86,655
3	-€ 350,223,082	-€ 9,359	53	€ 3,451,124,826	€ 92,052
4	-€ 131,686,069	-€ 3,453	54	€ 3,687,118,471	€ 97,857
5	-€ 27,354,601	-€ 699	55	€ 3,960,717,328	€ 104,098
6	-€ 2,031,340	-€ 58	56	€ 4,071,935,299	€ 109,807
7	€ 4,849,032	€ 133	57	€ 4,477,655,995	€ 115,830
8	€ 11,795,017	€ 308	58	€ 4,525,270,274	€ 122,590
9	€ 19,183,448	€ 520	59	€ 4,886,482,542	€ 129,428
10	€ 32,049,450	€ 843	60	€ 5,215,886,760	€ 136,376
11	€ 43,970,056	€ 1,168	61	€ 5,347,305,724	€ 143,778
12	€ 56,642,204	€ 1,501	62	€ 5,681,274,823	€ 151,277
13	€ 71,444,668	€ 1,866	63	€ 5,997,688,918	€ 158,165
14	€ 84,703,279	€ 2,284	64	€ 6,344,970,355	€ 165,242
15	€ 104,184,936	€ 2,795	65	€ 6,454,105,506	€ 172,193
16	€ 130,456,853	€ 3,375	66	€ 6,790,168,761	€ 178,711
17	€ 151,565,187	€ 4,044	67	€ 6,910,307,129	€ 184,625
18	€ 179,547,078	€ 4,639	68	€ 7,199,763,162	€ 190,476
19	€ 190,332,379	€ 5,206	69	€ 7,350,016,162	€ 197,594
20	€ 216,767,687	€ 5,774	70	€ 7,726,510,304	€ 204,984
21	€ 251,136,213	€ 6,467	71	€ 8,061,398,629	€ 212,472
22	€ 267,477,697	€ 7,269	72	€ 8,378,204,191	€ 221,006
23	€ 306,014,390	€ 8,116	73	€ 8,663,393,128	€ 229,111
24	€ 340,303,613	€ 8,982	74	€ 8,862,891,196	€ 236,805
25	€ 367,651,612	€ 9,886	75	€ 9,433,701,508	€ 246,102
26	€ 410,537,083	€ 10,715	76	€ 9,614,600,044	€ 255,810
27	€ 422,671,127	€ 11,420	77	€ 10,170,821,933	€ 267,503
28	€ 480,045,250	€ 12,315	78	€ 10,412,237,530	€ 281,618
29	€ 503,803,638	€ 13,434	79	€ 11,242,357,124	€ 292,988
30	€ 541,336,940	€ 14,436	80	€ 11,605,826,683	€ 305,194
31	€ 587,637,325	€ 15,622	81	€ 11,762,059,665	€ 317,001
32	€ 644,224,001	€ 16,935	82	€ 12,303,638,009	€ 328,401
33	€ 681,156,476	€ 18,480	83	€ 12,965,837,696	€ 344,048
34	€ 786,833,734	€ 20,290	84	€ 13,999,513,186	€ 366,031
35	€ 825,964,179	€ 22,198	85	€ 14,855,118,531	€ 387,317
36	€ 925,296,705	€ 24,157	86	€ 15,048,662,009	€ 410,967
37	€ 961,778,351	€ 25,999	87	€ 16,746,531,148	€ 435,064
38	€ 1,081,868,823	€ 28,200	88	€ 17,112,163,746	€ 461,864
39	€ 1,165,170,440	€ 31,006	89	€ 18,728,070,062	€ 491,832
40	€ 1,261,668,153	€ 33,648	90	€ 19,675,082,379	€ 525,777
41	€ 1,357,996,461	€ 36,145	91	€ 21,549,524,364	€ 565,643
42	€ 1,461,584,116	€ 38,590	92	€ 23,379,682,364	€ 620,432
43	€ 1,601,620,173	€ 41,927	93	€ 26,337,034,014	€ 691,303
44	€ 1,698,180,430	€ 45,274	94	€ 28,629,866,451	€ 768,496
45	€ 1,819,512,785	€ 48,591	95	€ 33,296,338,035	€ 868,675
46	€ 1,985,371,702	€ 52,469	96	€ 38,815,150,674	€ 1,041,491
47	€ 2,139,149,402	€ 57,336	97	€ 48,693,272,174	€ 1,297,201
48	€ 2,385,410,195	€ 62,082	98	€ 65,450,423,948	€ 1,712,739
49	€ 2,520,176,125	€ 67,274	99	€ 94,075,818,312	€ 2,524,137
50	€ 2,749,833,086	€ 73,442	100	€ 236,958,825,570	€ 6,380,234
		Gesamtvermögen:		€ 1,000,221,482,909	

Anmerkung: Eine Besonderheit von imputierten Datensätzen ist, dass die daraus gewonnenen Resultate immer nur Durchschnittswerte über alle Imputationen sein können. Nach Rubin und Little (2002) können solche Ergebnisse deshalb nicht als Grundlage für weitere Berechnungen verwendet werden. Weiters ist bei dieser detaillierten Darstellung Vorsicht geboten, weil die Basis für jedes Perzentil nur jeweils eine kleine Menge von Beobachtungen bildet und die Standardfehler daher hoch sind.

Anhang II: Schätzung der Verteilungsfunktion

Daten einlesen

Hier werden die entsprechenden Datenreihen (Nettovermögen und Gewichte) eingelesen.

```
SetDirectory[
  "/Users/joko/Dropbox/Vermögensverteilung/Stata und Daten/mathematica input"];

ω1 = Reverse[ReadList["wealth1.dat"]];
ω2 = Reverse[ReadList["wealth2.dat"]]; ω3 = Reverse[ReadList["wealth3.dat"]];
ω4 = Reverse[ReadList["wealth4.dat"]]; ω5 = Reverse[ReadList["wealth5.dat"]];
w1 = Reverse[ReadList["gewichte1.dat"]]; w2 = Reverse[ReadList["gewichte2.dat"]];
w3 = Reverse[ReadList["gewichte3.dat"]]; w4 = Reverse[ReadList["gewichte4.dat"]];
w5 = Reverse[ReadList["gewichte5.dat"]];
n = 30;
imp = 5;
```

Matrizen-Definition

Die Matrix Ω sammelt die Nettovermögen über alle Implicates.
Die Matrix Γ sammelt die Gewichte über alle Implicates (in % der Grundgesamtheit).
Die Matrizen P und X sammen die Ranks und Vermögenswerte der Perzentilgrenzen.
Die Matrizen Π, T und A sammeln die Pareto-Verteilungen (Π) bzw. deren Alphas (A) und die dazugehörigen P-Werte (T).

```
Ω = {ω1, ω2, ω3, ω4, ω5};
Γ = {w1 / Total[w1] * 100, w2 / Total[w2] * 100,
   w3 / Total[w3] * 100, w4 / Total[w1] * 100, w5 / Total[w5] * 100}
P = Table[Subscript[r, i, j], {i, 5}, {j, n}]; (*Matrix CutOffRanks*)
X = Table[Subscript[s, i, j], {i, 5}, {j, n}];(*Matrix CutOffValues*)
Π = Table[Subscript[t, i, j], {i, 5}, {j, n}];(*Matrix Pareto-Verteilungen*)
T = Table[Subscript[u, i, j], {i, 5}, {j, n}]; (*Matrix p-Werte*)
A = Table[Subscript[v, i, j], {i, 5}, {j, n}]; (*Matrix Alphas*)
```

Berechnung der Perzentilgrenzen

Im Folgenden werden die Perzentilgrenzen und die dazugehörigen Vermögenswerte berechnet und in die entsprechenden Matrizen eingetragen.

```
For[i = 1, i ≤ imp, i++,
  cumpw = Accumulate[Γ[[i]]];
  lcutoffnumber = {};
  For[j = 1, j ≤ n, j ++,
   length = Select[cumpw, # > j &];
   cutoffnumber = Length[cumpw] - Length[length];
   AppendTo[lcutoffnumber, cutoffnumber]; P[[i]] = lcutoffnumber];
  X[[i]] = Part[Ω[[i]], P[[i]]]];
P
X
```

```
{{21, 46, 71, 97, 123, 145, 169, 192, 215, 240, 261, 285, 313, 341, 361, 386,
  412, 438, 459, 488, 512, 535, 558, 584, 610, 638, 662, 685, 711, 738},
 {24, 47, 69, 97, 119, 145, 170, 196, 219, 239, 263, 287, 310, 335, 360, 383,
  408, 436, 460, 481, 508, 529, 558, 581, 603, 630, 652, 678, 704, 730},
 {25, 50, 75, 98, 122, 144, 172, 194, 215, 239, 262, 285, 312, 337, 361, 384,
  409, 436, 465, 485, 514, 541, 565, 591, 617, 641, 666, 691, 718, 741},
 {26, 52, 76, 102, 124, 148, 173, 195, 218, 239, 263, 285, 311, 339, 364,
  388, 410, 439, 465, 486, 510, 538, 562, 584, 612, 635, 662, 685, 712, 737},
 {22, 47, 74, 98, 124, 148, 173, 195, 216, 241, 265, 292, 315, 340, 362, 383,
  411, 437, 463, 486, 515, 538, 562, 582, 607, 637, 660, 688, 716, 741}}
```

```
{{2 771 220, 1 940 431, 1 491 249, 1 194 820, 933 582, 810 549,
  725 027, 653 500, 578 266, 533 620, 500 600, 467 798, 435 744, 414 614,
  387 020, 365 706, 341 410, 326 894, 316 000, 303 988, 290 955, 281 242,
  268 122, 257 630, 247 423, 234 959, 226 224, 220 000, 212 159, 204 433},
 {4 599 361, 2 806 400, 1 782 401, 1 236 753, 1 056 000, 822 000, 762 240,
  674 595, 605 880, 557 663, 521 027, 481 400, 456 633, 431 911, 404 156,
  380 909, 355 600, 332 435, 322 548, 311 264, 298 934, 287 809, 269 828,
  261 466, 250 200, 241 753, 233 801, 225 200, 215 880, 206 882},
 {2 843 000, 1 672 400, 1 315 775, 1 098 300, 900 402, 796 899, 702 932,
  651 087, 585 112, 538 225, 504 200, 476 459, 455 424, 424 142, 403 792,
  382 200, 364 206, 344 538, 328 772, 317 432, 304 135, 289 811, 276 837,
  262 000, 250 241, 241 658, 234 165, 226 437, 217 819, 212 561},
 {3 442 834, 2 028 002, 1 455 544, 1 106 411, 948 034, 840 332, 746 500,
  680 346, 605 331, 556 078, 519 763, 497 127, 470 775, 432 910, 410 027,
  385 265, 366 004, 339 241, 326 884, 315 300, 306 054, 293 161, 281 500,
  265 371, 255 530, 248 106, 235 507, 229 623, 220 000, 211 371},
 {2 719 144, 1 787 689, 1 290 388, 1 057 280, 884 000, 782 250, 696 002,
  628 009, 581 729, 531 500, 499 415, 462 890, 431 164, 411 941, 387 020,
  369 033, 347 333, 334 188, 319 826, 309 800, 296 344, 288 422, 274 025,
  263 469, 251 900, 241 228, 233 800, 225 496, 215 481, 208 003}}
```

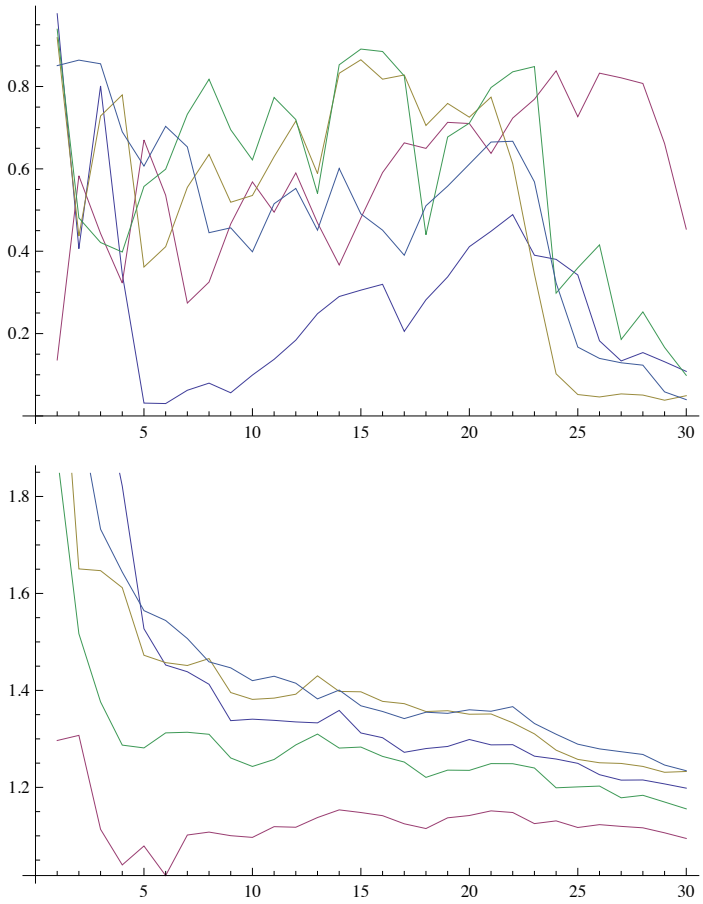
Pareto-Schätzung und Cramer-von-Mises-Test

Schätzung der Verteilung und Durchführung des Tests

Hier werden die entsprechenden Paretoverteilungen für alle Perzentilgrenzen über alle Implicates geschätzt und die statistische Plausibilität dieser Schätzung mittels eines Cramer von Mises Tests überprüft. Abschließend werden die Teststatistiken und geschätzten Alpha-Parameter der jeweiligen Verteilungen über die letzten dreißig Perzentile abgebildet.

```
For[i = 1, i ≤ imp, i++,
 For[j = 1, j ≤ n, j++,
  Π[[i, j]] = EstimatedDistribution[
    Select[Ω[[i]], # ≥ X[[i, j]] &], ParetoDistribution[X[[i, j]], α]];
  A[[i, j]] = Replace[α, FindDistributionParameters[
     Select[Ω[[i]], # ≥ X[[i, j]] &], ParetoDistribution[X[[i, j]], α]]];
  T[[i, j]] = DistributionFitTest[Select[Ω[[i]], # ≥ X[[i, j]] &],
    Π[[i, j]], "CramerVonMises"]]]
Π;
T;
A;

ListLinePlot[T]
ListLinePlot[A]
```



Minima und Maxima der p-Werte ausheben

An dieser Stelle wird die Struktur der p-Werte genauer analysiert. Im Detail werden vor allem die Minima und Maxima der p-Werte im jeweiligen Perzentil (d.h. über alle Implicates) verglichen. Als Ansatzpunkt für die Pareto-Verteilung wurde in Folge jener Punkt gewählt an dem die Minima der p-Werte über alle Imputationen maximal sind. Anders ausgedrückt wurde versucht die Absicherung gegen eine Fehleinschätzung hinsichtlich der Verteilungsannahmen zu maximieren.

```
minis = {};
maxis = {};
For[j = 1, j ≤ n, j++,
 mini = Min[{T[[1, j]], T[[2, j]], T[[3, j]], T[[4, j]], T[[5, j]]}];
 AppendTo[minis, mini]]
For[j = 1, j ≤ n, j++,
 maxi = Max[{T[[1, j]], T[[2, j]], T[[3, j]], T[[4, j]], T[[5, j]]}];
 AppendTo[maxis, maxi]]
minis
maxis
Max[minis]
Position[minis, Max[minis]]
Max[maxis]
Position[maxis, Max[maxis]]
```

```
{0.135926, 0.406139, 0.421202, 0.322666, 0.0312194, 0.0300967, 0.0623624,
 0.0796877, 0.0562547, 0.0992766, 0.137477, 0.184129, 0.248271, 0.290001,
 0.305518, 0.319646, 0.205402, 0.282096, 0.33766, 0.411014, 0.448505, 0.488949,
 0.34461, 0.102294, 0.052, 0.046111, 0.0535458, 0.0507137, 0.0382005, 0.0395955}
```

```
{0.976766, 0.864081, 0.855182, 0.779636, 0.670182, 0.70327, 0.732606, 0.817599,
 0.694986, 0.621758, 0.773386, 0.719785, 0.588499, 0.852709, 0.890932,
 0.884981, 0.827959, 0.705429, 0.75887, 0.725475, 0.797514, 0.835682, 0.848367,
 0.837956, 0.726734, 0.832408, 0.821041, 0.80735, 0.661539, 0.453699}
```

```
0.488949
```

```
{{22}}
```

```
0.976766
```

```
{{1}}
```

Anhang III: Perzentilliste auf Basis der korrigierten Daten

Perz-entil	Gesamtvermögen im Perzentil	Durchschnittsvermögen im Perzentil	Perz-entil	Gesamtvermögen im Perzentil	Durchschnittsvermögen im Perzentil
1	-€ 5,491,830,186	-€ 143,347	51	€ 2,997,670,620	€ 80,687
2	-€ 851,460,999	-€ 22,396	52	€ 3,307,328,626	€ 87,161
3	-€ 352,235,977	-€ 9,314	53	€ 3,466,137,551	€ 92,465
4	-€ 130,321,132	-€ 3,390	54	€ 3,709,728,015	€ 98,378
5	-€ 25,867,881	-€ 680	55	€ 3,964,282,913	€ 104,570
6	-€ 2,028,056	-€ 57	56	€ 4,150,303,848	€ 110,344
7	€ 4,888,335	€ 133	57	€ 4,422,465,916	€ 116,365
8	€ 12,150,970	€ 313	58	€ 4,612,697,208	€ 123,169
9	€ 18,948,024	€ 525	59	€ 4,991,646,628	€ 130,145
10	€ 32,382,112	€ 847	60	€ 5,087,187,943	€ 137,021
11	€ 43,200,756	€ 1,171	61	€ 5,540,623,468	€ 144,687
12	€ 57,419,467	€ 1,503	62	€ 5,716,775,305	€ 152,037
13	€ 73,320,760	€ 1,876	63	€ 5,935,207,031	€ 158,946
14	€ 84,210,376	€ 2,299	64	€ 6,328,856,180	€ 166,007
15	€ 104,778,599	€ 2,809	65	€ 6,395,678,723	€ 172,763
16	€ 128,453,312	€ 3,387	66	€ 6,935,411,750	€ 179,399
17	€ 153,723,229	€ 4,053	67	€ 7,038,328,270	€ 185,416
18	€ 179,811,797	€ 4,656	68	€ 7,156,248,769	€ 191,231
19	€ 193,071,343	€ 5,222	69	€ 7,402,561,055	€ 198,481
20	€ 217,996,343	€ 5,798	70	€ 7,680,458,821	€ 205,859
21	€ 245,306,962	€ 6,484	71	€ 8,125,472,192	€ 213,315
22	€ 269,627,869	€ 7,282	72	€ 8,391,578,758	€ 221,995
23	€ 308,107,437	€ 8,134	73	€ 8,596,868,259	€ 229,863
24	€ 346,012,313	€ 9,013	74	€ 9,070,948,198	€ 237,898
25	€ 372,477,084	€ 9,925	75	€ 9,318,691,747	€ 247,391
26	€ 403,367,966	€ 10,743	76	€ 9,804,058,985	€ 257,125
27	€ 437,570,617	€ 11,465	77	€ 10,123,657,942	€ 269,045
28	€ 464,772,081	€ 12,363	78	€ 10,640,965,751	€ 283,258
29	€ 516,491,736	€ 13,472	79	€ 11,152,696,422	€ 294,478
30	€ 538,392,030	€ 14,499	80	€ 11,362,715,108	€ 306,795
31	€ 600,387,947	€ 15,675	81	€ 12,269,421,672	€ 318,384
32	€ 635,033,538	€ 17,027	82	€ 12,296,480,811	€ 330,197
33	€ 701,434,012	€ 18,585	83	€ 13,087,706,189	€ 346,528
34	€ 764,497,336	€ 20,397	84	€ 14,012,243,821	€ 368,979
35	€ 830,947,524	€ 22,292	85	€ 14,875,455,727	€ 389,993
36	€ 922,964,935	€ 24,230	86	€ 15,385,765,867	€ 414,369
37	€ 980,952,230	€ 26,102	87	€ 16,746,754,331	€ 438,963
38	€ 1,068,212,515	€ 28,307	88	€ 17,764,694,616	€ 466,308
39	€ 1,190,504,553	€ 31,167	89	€ 18,347,820,078	€ 496,495
40	€ 1,296,701,590	€ 33,832	90	€ 20,251,800,140	€ 531,114
41	€ 1,347,223,403	€ 36,323	91	€ 21,757,091,796	€ 573,303
42	€ 1,479,940,426	€ 38,838	92	€ 23,635,883,047	€ 630,423
43	€ 1,577,039,431	€ 42,136	93	€ 26,349,802,682	€ 702,219
44	€ 1,695,678,901	€ 45,501	94	€ 30,213,151,014	€ 784,894
45	€ 1,860,447,265	€ 48,846	95	€ 33,089,917,775	€ 888,956
46	€ 1,979,027,923	€ 52,802	96	€ 40,397,367,382	€ 1,073,117
47	€ 2,205,959,092	€ 57,735	97	€ 50,791,826,986	€ 1,341,788
48	€ 2,343,784,250	€ 62,467	98	€ 66,601,914,082	€ 1,768,638
49	€ 2,563,660,981	€ 67,721	99	€ 101,014,993,783	€ 2,690,588
50	€ 2,826,810,210	€ 74,039	100	€ 469,058,597,640	€ 12,670,045
		Gesamtvermögen:		€ 1,248,599,886,785	

Anmerkung: Eine Besonderheit von imputierten Datensätzen ist, dass die daraus gewonnenen Resultate immer nur Durchschnittswerte über alle Imputationen sein können. Nach Rubin und Little (2002) können solche Ergebnisse deshalb nicht als Grundlage für weitere Berechnungen verwendet werden. Weiters ist bei dieser detaillierten Darstellung Vorsicht geboten, weil die Basis für jedes Perzentil nur jeweils eine kleine Menge von Beobachtungen bildet und die Standardfehler daher hoch sind.

Anhang IV: Samplekorrektur

Daten einlesen

Hier werden die wesentlichen Start-Daten aus dem HFCS eingelesen (Nettovermögen und originale Gewichte über alle 5 Imputationen). Die für die Berechnung der Ausweicheffekte relevanten Daten (Vermögenskomponenten) erfolgt später.

```
SetDirectory["YOUR_DIRECTORY_HERE"];

ω1 = Reverse[ReadList["wealth1.dat"]];
ω2 = Reverse[ReadList["wealth2.dat"]]; ω3 = Reverse[ReadList["wealth3.dat"]];
ω4 = Reverse[ReadList["wealth4.dat"]]; ω5 = Reverse[ReadList["wealth5.dat"]];
w1 = Reverse[ReadList["kgewichte1.dat"]];
w2 = Reverse[ReadList["kgewichte2.dat"]];
w3 = Reverse[ReadList["kgewichte3.dat"]];
w4 = Reverse[ReadList["kgewichte4.dat"]];
w5 = Reverse[ReadList["kgewichte5.dat"]];
```

Definition relevanter Matrizen und Listen

Nettovermögen und Haushaltsgewichte

Ω bezeichnet die Matrix der Nettovermögen über alle 5 Implicates.
Γ bildet die originalen Gewichte aus dem HFCS in einer Matrix über alle 5 Implicates ab.
Λ sammelt schließlich alle Modifikationen der Gewichte in einer neuen Gewichtsmatrix, $\Lambda\Lambda$ repräsentiert dieselbe Liste ergänzt um die Gewichtungen der neu generierten Haushalte.

```
Ω = {ω1, ω2, ω3, ω4, ω5};
Γ = {w1, w2, w3, w4, w5};
Λ = {};
ΛΛ = {};
```

Listen zur Erfassung der 5 Implicates

Die Liste Δ dient der Erfassung jener Beobachtungen, die am oberen Rand der Verteilung entfernt und durch Schätzwerte aus der Verteilungsstatistik ersetzt werden. Die Liste $\Delta\Delta$ zeigt das kumulierte Gewicht dieser Haushalte an.
Analog dazu zeigt die Liste H jene Beobachtungen an, die über dem für die Pareto-Verteilung relevanten CutOff-Wert (γ_{1-5} – siehe Kapitel "Parameterdefinition") liegen, und die Liste HH das kumulierte Gewicht dieser Haushalte.
Die Liste N sammelt die Zahl neu zu generierender Haushalte für alle 5 Implicates.
Die Liste NN ist im Wesentlichen die Differenz der Listen N (neu zu generierende Haushalte) und Δ (zu eliminierende Haushalte), dividiert durch die Gesamtzahl der Population (korrigiert um die eliminierten Haushalte). Da sich aus dieser Kalkulation jener Faktor ergibt, um den die verbleibenden originalen Gewichte abgeschmolzen/aufgewertet werden müssen um die Population wieder vollständig abzubilden, kann man von NN auch als “Abschmelzungs-“ bzw. “Aufwertungsfaktor”

sprechen.
Die Listen K und KK dienen der Fehlerkontrolle nach der Anpassung der Gewichte.
Die Listen υ_{1-5} sammeln die zufällig zur Fehlerkorrektur neu gezogenen Haushalte.

```
Δ = {};
ΔΔ = {};
H = {};
HH = {};
N = {};
NN = {};
K = {};
KK = {};
υ1 = {}; υ2 = {}; υ3 = {}; υ4 = {}; υ5 = {};
```

Parameterdefinition

Verteilungsparameter

α_{1-5} bezeichnen die geschätzten Parameter (Paretos Alpha) über alle 5 Implicates.
γ_{1-5} bezeichnen die geschätzten Minimalwerte, ab denen die Daten gemäß der Pareto-Verteilung verteilt sind, für alle 5 Implicates.
χ_{1-5} bezeichnen jenes Vermögensniveau, ab denen nicht mehr auf die Daten des HFCS, sondern auf die daraus errechnete Pareto-Verteilung zurückgegriffen wird.
ρ_{1-5} bezeichnen den Prozentanteil jener Haushalte mit einem Nettovermögen größer als χ aus der Menge all jener Haushalte, die gemäß einer Pareto-Verteilung verteilt sind (d.h. ein Nettovermögen größer als γ aufweisen), für alle 5 Implicates.

```
α1 = 1.28808; α2 = 1.14815; α3 = 1.3332; α4 = 1.24881; α5 = 1.36649;
γ1 = 281 242; γ2 = 287 809; γ3 = 289 811; γ4 = 293 161; γ5 = 288 422;
χ1 = 4 000 000; χ2 = 4 000 000; χ3 = 4 000 000; χ4 = 4 000 000; χ5 = 4 000 000;
ρ1 = 1 - CDF[ParetoDistribution[γ1, α1], χ1];
ρ2 = 1 - CDF[ParetoDistribution[γ2, α2], χ2];
ρ3 = 1 - CDF[ParetoDistribution[γ3, α3], χ3];
ρ4 = 1 - CDF[ParetoDistribution[γ4, α4], χ4];
ρ5 = 1 - CDF[ParetoDistribution[γ5, α5], χ5];
```

Sonstige Variablen

```
imp = 5;
```

Datenbereinigung und -aufbereitung

Berechnung der Zahl Haushalte mit Nettovermögen > χ

Im Folgenden werden jene Haushalte bestimmt, deren Nettovermögen größer χ ist. Der Output am Ende dieses Subkapitels zeigt an wieviele Beobachtungen aus dem jeweiligen Implicate gestrichen werden und wievielen Haushalten dies unter Berücksichtigung der Gewichtung entspricht.

```
For[i = 1, i ≤ imp, i++, δ = Length[Ω[[i]]] - Length[Select[Ω[[i]], # < χᵢ &]];
  AppendTo[Δ, δ]; AppendTo[ΔΔ, Total[Take[Γ[[i]], δ]]]];
Δ
ΔΔ
```

```
{8, 30, 12, 19, 11}
```

```
{11 374.2, 44 079.1, 16 564.3, 28 225.3, 16 975.9}
```

Berechnung der Zahl der Haushalte mit Nettovermögen > γ und < χ

Im Folgenden wird die Zahl der Haushalte am oberen Ende der Vermögensverteilung berechnet, deren Verteilung die zuvor geschätzte Pareto-Verteilungsfunktion beschreibt. Die beiden am Ende dieses Subkapitels angezeigten Daten entsprechen wiederum der Zahl der für die Bestimmung der Pareto-Verteilung relevanten Beobachtungen bzw. Haushalte pro Implicate.

```
For[i = 1, i ≤ imp, i++, h = Length[Select[Ω[[i]], χ₁ ≥ # ≥ γᵢ &]]; AppendTo[H, h]]
For[i = 1, i ≤ imp, i++,
 hh = Total[Drop[Drop[Γ[[i]], Δ[[i]]], -Length[Select[Ω[[i]], # < γᵢ &]]]];
 AppendTo[HH, hh]]
H
HH
```

```
{527, 499, 529, 519, 527}
```

```
{817 418., 785 924., 813 359., 801 910., 812 560.}
```

Berechnung der Zahl neu zu generierender Haushalte

Die in der Liste HH abgebildeten Werte geben an, wieviele Haushalte sich zwischen einem Vermögenswert von γ und einem von χ befinden. In Kombination mit ρ (ρ gibt jenen Teil der Haushalte, der ein Vermögen > χ aufweist, als Teilmenge all jener Haushalte, die ein Vermögen von > γ aufweisen, an) lässt sich nun errechnen wieviele Haushalte einen solchen Vermögenswert > χ aufweisen. Es ergibt sich hier also die Summe jener Haushalte, die dem Datensatz zur Korrektur des unterstellten Bias hinzugefügt werden muss.

```
For[i = 1, i ≤ imp, i++, AppendTo[N, Round[HH[[i]] * ρᵢ / (1 - ρᵢ)]]]
N
```

```
{27 654, 40 252, 25 341, 31 895, 22 982}
```

(Lineare) Abschmelzung/Aufwertung der verbleibenden Haushalte

An dieser Stelle wird zuerst die Zahl jener Haushalte berechnet, die abgeschmolzen werden müssen. Diese ergibt sich aus der Differenz der hinzugefügten Haushalte (N) minus der Zahl eliminierter Haushalte (ΔΔ). Ein Vergleich dieser Werte für alle fünf Implicates zeigt, dass dies für die Implicates 1, 3 bis 5 eine Abschmelzung bedeutet, für das zweite Implicate allerdings eine Aufwertung der verbliebenen Haushalte impliziert (die Zahl der eliminierten Haushalte ist im zweiten Implicate größer als jene der hinzugefügten). Der jeweils errechnete Wert wird in Folge durch einen Bezug zur Grundgesamtheit normalisiert (NN) - die Abschmelzung/Aufwertung erfolgt also proportional zu den originalen Gewichtungen. Die Matrix Λ sammelt schließlich die abgeschmolzenen Gewichte der nicht-ausgeschiedenen Haushalte.

```
For[i = 1, i ≤ imp, i++, AppendTo[NN, (N[[i]] - ΔΔ[[i]])/Total[Γ[[i]]] - ΔΔ[[i]]];
 AppendTo[Λ, Drop[Γ[[i]] * (1 - NN[[i]]), Δ[[i]]]]]
NN
Λ;
```

```
{0.00432677, -0.00102607, 0.00233586, 0.000979709, 0.00159866}
```

Generierung der Gewichte der neuen Haushalte

Im Folgenden werden die Gewichte für die neu generierten Haushalte dem Sample der Gewichte hinzugefügt (die neu zu ziehenden Haushalte weisen alle ein Gewicht von 1 auf).

```
For[i = 1, i ≤ imp, i++, AppendTo[ΛΛ, Join[Table[1, {ii, N[[i]]}], Λ[[i]]]]]

ΛΛ;
```

Kontrollrechnungen: Validierung der bisherigen Ergebnisse durch Kontrolle der Summen

Hier werden die Längen und die Summen der Gewichtsvektoren durch Differenzenbildung kontrolliert um die Gewichtsvektorskonstruktion zu überprüfen.

```
For[i = 1, i ≤ imp, i++, check = Total[Γ[[i]]] - Total[ΛΛ[[i]]]; AppendTo[K, check]]
For[i = 1, i ≤ imp, i++,
 k = Length[ΛΛ[[i]]] - (Length[Λ[[i]]] + N[[i]]); AppendTo[KK, k]]
K
KK
```

```
{0., 4.65661 × 10^-10, 4.65661 × 10^-10, 0., 4.65661 × 10^-10}
```

```
{0, 0, 0, 0, 0}
```

Zufallsziehung der neuen Haushalte

Zufallsziehung der neuen Haushalte

Im vorliegenden Abschnitt werden die Haushalte zuerst für jedes Implicate gezogen, dann sortiert und zu den neuen Vektoren ϕ_{1-5} zusammengefasst. Aufgrund der unterschiedlich Zahl an zu ergänzenden Haushalten haben diese Vektoren nun (ebenso wie die oben erstellten Gewichtsvektoren) eine unterschiedliche Länge. Die Summe der Gewichte ist aber für jedes Implicate gleich (es wird also immer dieselbe Grundgesamtheit repäsentiert). Hier sind verschiedene Optionen möglich: So kann die Zahl der Zufallsziehungen pro Implicate erhöht werden (n), wobei dann ein gemittelter Durchschnitt über die Zahl an Ziehungen eingetragen wird. Diese Option ist geeignet die Streuweite der Ergebnisse signifikant zu reduzieren und damit eine statistisch gesehen präzisere Schätzung zu liefern (Allerdings ist hier mit einem erhöhten Rechenaufwand zu kalkulieren: Bei n = 1 und 8 GB RAM dauert eine Ziehung zwischen 3 und 5 Minuten, die sich für n > 1 mulitplizieren). Weiters besteht die Möglichkeit aus kalkulatorischer Vorsicht das maximal gezogene Vermögen (max) zu beschränken. Die beiden Rechenzeilen erledigen die notwendigen Berechnungen und Umformungen.

```
n = 5;
max = 1 000 000 000;
ϒ1 = {}; ϒ2 = {}; ϒ3 = {}; ϒ4 = {}; ϒ5 = {};

For[i = 1, i ≤ imp, i++, For[k = 1, k ≤ n, k++, For[j = 1,
    Length[υi] < N[[i]], j++, z = RandomVariate[ParetoDistribution[γi, αi]];
    If[max > z > χi, AppendTo[υi, z], 0]];
  AppendTo[ϒi, Sort[υi, Greater]]; υi = {}]; φi = Mean[ϒi]]
For[i = 1, i ≤ imp, i++, ϕi = Join[φi, Drop[Ω[[i]], Δ[[i]]]]]
```

Kontrollrechnungen

Hier werden drei Kontrollrechnungen durchgeführt : (1) Längenvergleich von Datenvektor und Gewichtsvektor für jedes Implicate, (2) Längenvergleich von Datenvektor und Zahl der gezogenen Haushalte (unter Berücksichtigung der zuerst ausgeschiedenen), (3) Summenvergleich über alle Gewichtsvektoren.

```
0 == Length[ϕ1] - Length[ΛΛ[[1]]]
0 == Length[ϕ2] - Length[ΛΛ[[2]]]
0 == Length[ϕ3] - Length[ΛΛ[[3]]]
0 == Length[ϕ4] - Length[ΛΛ[[4]]]
0 == Length[ϕ5] - Length[ΛΛ[[5]]]

2380 == Length[ϕ1] - Length[φ1] + Δ[[1]]
2380 == Length[ϕ2] - Length[φ2] + Δ[[2]]
2380 == Length[ϕ3] - Length[φ3] + Δ[[3]]
2380 == Length[ϕ4] - Length[φ4] + Δ[[4]]
2380 == Length[ϕ5] - Length[φ5] + Δ[[5]]

0 == Round[Total[ΛΛ[[1]]] - Total[w1]]
0 == Round[Total[ΛΛ[[2]]] - Total[w2]]
0 == Round[Total[ΛΛ[[3]]] - Total[w3]]
0 == Round[Total[ΛΛ[[4]]] - Total[w4]]
0 == Round[Total[ΛΛ[[5]]] - Total[w5]]
True

True

True

True

True

True

True

True

True

True

True

True

True

True

True

True
```