



Längsschnittanalysen mit EU-SILC

Endbericht

August Gächter



Bericht an das International Centre for Migration Policy Development (ICMPD).

Inhalt

Zusammenfassung	5
1. Problemstellung	7
1.1. Inhaltliche Frage.....	7
1.2. Analytische Frage.....	7
2. Methoden der Panelanalyse	9
2.1. Kausalanalyse	9
2.2. Stichprobeneinheit.....	11
2.3. Fixed Effects.....	13
2.4. Mehrebenenanalyse.....	14
3. Stichprobe	16
3.1. Stichprobenschwund	16
3.2. Bestimmungsgründe des Schwundes	17
4. Deskriptive Ergebnisse	17
4.1. Das äquivalisierte Nettohaushaltseinkommen pro Jahr	17
4.2. Quellen des Wachstums des äquivalisierten Haushaltseinkommens	26
4.3. Quellen des Wachstums des Haushaltseinkommens	29
4.4. Fazit: Konvergenz?	31
5. Panelanalyse.....	33
5.1. Fragestellung.....	33
5.2. Die Erwerbsintensität des Haushalts.....	33
5.3. Auf die Erwerbsintensität abgeglichenes Nettohaushaltseinkommen	35
5.4. Stichprobeneigenschaften	37
5.5. Wachstumskurvenanalyse	39
5.6. Fazit.....	46
6. Quellenangaben.....	52



Zusammenfassung

Daten, bei denen dieselben Personen im Zeitverlauf mehrmals über dieselben Dinge befragt werden, haben den Vorteil, statistisch robuste Aussagen über Ursache und Wirkung zuzulassen, was bei einmaligen Befragungen nicht möglich ist. Solche Mehrfachbefragungen derselben Personen – sogenannte Panelbefragungen – gibt es in Österreich wenige. Eine ist EU-SILC, die jährlich durchgeführt wird, und bei der die Haushalte mit all ihren Mitgliedern vier Mal befragt werden, so dass zwischen der ersten und der letzten Befragung drei Jahre liegen. In der vorliegenden Untersuchung wurden die zwischen 2008 und 2010 erstmals und zwischen 2011 und 2013 letztmalig befragten Haushalte ausgewertet. Die Forschungsfrage war zweifach, nämlich ob die Fallzahlen von EU-SILC ausreichen, um spezifisch für Paneldaten entwickelte Analysemethoden erfolgreich anwenden zu können, und ob sich auf diese Weise ergründen lässt, in welchem Maß und aufgrund welcher Ursachen die jährlichen Haushaltseinkommen von eingewanderten und nicht eingewanderten Haushalten sich im Zeitverlauf einander annähern oder es nicht tun.

Es wurde zwischen Haushalten aus Österreich, aus den EU15/EFTA Staaten und aus allen übrigen Staaten unterschieden. Etwas vereinfachend kann man sagen, die Verlaufsmuster der Haushaltseinkommen über die vier Befragungen seien im Wesentlichen zwischen diesen drei Gruppen von Haushalten parallel (siehe Abb. 1 und 2). Das bedeutet aber, dass die von außerhalb der EU15/EFTA Staaten zugezogenen Haushalte gegenüber denjenigen aus Österreich nicht aufholen, oder wenn, dann nur unmerklich langsam. Eine Annäherung (Konvergenz) kann sich demnach nur im Generationswechsel ereignen. Man muss aber deutlich darauf hinweisen, dass dieser Befund sich auf die Zeit von 2008 bis 2013 bezieht, und dass vier Messzeitpunkte vielleicht nicht genug sind, um in die Zukunft zu extrapolieren.

Um die Ursachen dieser Parallelität aufzuklären, wurden panelstatistische Methoden (Mehrebenenanalyse) eingesetzt. Sie erweisen sich dabei als grundsätzlich anwendbar, wobei man aber bedingt durch die Fallzahlen auf Grenzen stößt. Besonders bei der relativ kleinen Gruppe der Haushalte aus EU15/EFTA Staaten treten Schwierigkeiten auf. Steigern lassen sich die Fallzahlen, indem man weitere Jahrgänge der SILC-Daten abwartet. Eine Verlängerung des Panels von vier auf sechs Befragungen wäre ebenfalls hilfreich, nicht nur wegen der Fallzahlen, sondern auch wegen der Extrapolierbarkeit in die Zukunft.

Mehrere Ursachen für die Parallelität werden erkennbar. Zugezogene Haushalte ziehen nur aus steigender Erwerbsbeteiligung der Haushaltsmitglieder und aus Tätigkeitswechseln Gewinn, während sie von Bildungszuwächsen im Haushalt keinen einkommensbezogenen Nutzen haben, jedenfalls nicht kurzfristig. Haushalte aus Österreich dagegen profitieren nicht nur aus größerer Erwerbsbeteiligung und aus Tätigkeitswechseln, sondern gewinnen auch durch Bildungszuwächse an Einkommen dazu. Bei ihnen gibt es außerdem noch weitere Zuwachsgründe, die nicht ermittelt werden konnten, die bei den Haushalten von außerhalb der EU15/EFTA Staaten aber nachweislich fehlen.

1. Problemstellung

1.1. Inhaltliche Frage

Wir möchten den möglichst genau modellierten Trendverlauf des äquivalisierten, d.h. bedarfsgewichteten Nettojahreseinkommens der Haushalte über den Zeitraum ihrer jeweils vier Befragungen von zwei Gruppen von Haushalten, nämlich den eingewanderten und den nicht eingewanderten. Wir wollen wissen, ob die Trends konvergieren, wie stark und warum.

1.2. Analytische Frage

Die Entwicklung des durchschnittlichen Einkommens einer Bevölkerung hängt zum einen von den Einkommenszuwächsen der Bezieherinnen und Bezieher von Einkommen ab, zum anderen aber auch von den Einkommen der neu hinzukommenden und der abgehenden Einkommensbezieherinnen und Einkommensbezieher. Wenn in einer Bevölkerung, zum Beispiel, Jugendliche einen wachsenden Anteil an den Einkommensbezieherinnen und Einkommensbeziehern bestreiten, dann kann dadurch das Durchschnittseinkommen sinken, selbst wenn von niemandem das persönliche Einkommen sinkt. Genau gleich kann sich das mit Einwanderung ereignen. Das Durchschnittseinkommen einer Bevölkerung mit dem Geburtsstaat A könnte rapide zurückgehen, wenn eine bestehende Bevölkerung, die im Wesentlichen um die 50 Jahre alt und etabliert ist, eine größere Zahl von 20 bis 25 jährigen neu Zuziehenden hinzukommt, die wesentlich weniger verdienen. Um falsche Schlüsse zu vermeiden, ist es in einer solchen Situation wichtig, die Einkommensentwicklung der etablierten Bevölkerung separat von jener der neu hinzukommenden Bevölkerung zu betrachten. Das ist in den üblichen Datenquellen zum Einkommen meist nicht möglich, weil sie keine Angabe zum Zeitpunkt des Aufenthaltsbeginns enthalten und nicht einmal zum Geburtsort, sondern bestenfalls zur Staatsangehörigkeit.

Selbst wenn man eine derartige Unterscheidung zu treffen in der Lage ist, besteht noch immer die Gefahr von Fehlschlüssen. Es könnte, etwa, sein, dass man bei der etablierten Bevölkerung aus dem Staat A ab einem bestimmten Zeitpunkt einen besonders starken Zuwachs der Einkommen und eine besonders rasche Annäherung an die Einkommen der inländischen Haushalte beobachtet. Wenn diese etablierte Bevölkerung um die 50 Jahre alt ist, dann werden neu erworbene Bildungsabschlüsse eher nicht die Ursache für die Änderung des Trends bei den Einkommen in Frage kommen. Die Ursache könnte aber sein, dass etablierte Haushalte mit unterdurchschnittlichen

Einkommenszuwächsen wegziehen und nur jene zurücklassen, die in Bezug auf das Einkommen besonders erfolgreich sind. Um keine falschen Schlüsse zu ziehen, wird man sich informieren müssen, wie sich die Einkommen der nicht wegziehenden Haushalte entwickelt haben, und wird dann vielleicht feststellen, dass ihre Einkommensentwicklung sich nicht verändert hat, sondern vor dem bestimmten Zeitpunkt ebenso günstig war wie danach.

Alle diese Vorkehrungen treffen sich von selbst, wenn man über einen längeren Zeitraum immer wieder dieselben Haushalte befragen kann. Dadurch gelingt es festzustellen, wie sich ein Durchschnittseinkommen unbeeinflusst von Zu- und Abgängen entwickelt. Solche Panelbefragungen gehen aber noch weit tiefer in ihrem Vermögen, äußere Einflüsse aus einem Durchschnitt draußen zu halten. Ihre Durchführung wird oftmals damit begründet, dass dadurch eine Vielzahl von Voraussetzungen, die man nie alle messen könnte, weil der Aufwand jeden für eine Befragung akzeptablen oder möglichen Aufwand sprengen würde, gar nicht gemessen werden müssen, weil sie bezogen auf die jeweilige Person oder den jeweiligen Haushalt ohnehin unveränderlich sind. Dabei denkt man etwa an die ganze genetische Ausstattung ebenso wie an sozial erworbene Arten, mit ihr und mit der Umwelt umzugehen. Die Veränderlichkeit der Einkommen muss und kann in einer Panelbefragung nicht auf die Veränderlichkeit von Merkmalen zurückgeführt werden, die in der einzelnen, erwachsenen Person nicht (mehr) veränderlich sind oder nur mehr in einem Ausmaß, von dem keine merklichen Auswirkungen zu erwarten sind. In Befragungswellen mit immer wechselnder Stichprobe würde man sich darauf nicht verlassen können oder dürfen.

Gerade aber wenn die Einkommensverläufe einer eingewanderten Bevölkerung nachgezeichnet werden sollen, besteht der zuletzt beschriebene Vorteil von Panelbefragungen nur in eingeschränkter Weise. Wir wissen viel zu wenig über die Anpassungsleistungen, die Einwanderinnen und Einwanderer in der kurzen, mittleren und längeren Frist erbringen, um davon ausgehen zu können, dass bei ihnen die Unveränderlichkeiten dieselben seien wie bei der ortsfesten Bevölkerung, mit der wir ihre Einkommensverläufe vergleichen. Oberflächlich sichtbar und hörbar wird das im sukzessiven Gebrauch der dominanten Verhaltensweisen und der Sprache des Orts, auf den sie zugezogen sind, aber allein das in einer Panelbefragung adäquat abzubilden, wie es sicherlich der Fall sein müsste, ist eine große Herausforderung. Gerade mit Bezug auf eine eingewanderte Bevölkerung, speziell auch eine erst vor kurzem eingewanderte Bevölkerung, hat eine Panelbefragung, die darauf nicht speziell ausgerichtet ist, manche der üblichen Vorteile daher nur in verringertem Maß. Die Erwartung wird daher sein müssen, dass sie die Einkommensverläufe von eingewanderten Bevölkerungsteilen zwar ebenso gut abbilden, aber weniger gut erklären kann, wie jene einer nicht migrierten Bevölkerung.

Wenig sinnvoll wäre es, eine Debatte zu führen, welche Betrachtungsweise der Einkommensverläufe die richtige sei. Es kommt auf den Zweck an. Die Panelbetrachtung liefert das Bild, wie es sich für den einzelnen Haushalt selbst darstellt, der Zeitverlauf liefert das Bild von außen, also wie der einzelne Haushalt den Rest der Gesellschaft sieht. Der Unterschied zwischen den beiden Bildern kann frappant sein, und von da her stammt vielleicht der Befund aus der Jugendforschung, dass die eigene Zukunft rosig gesehen wird, jene der Gesellschaft aber als schleichende Katastrophe.

2. Methoden der Panelanalyse

2.1. Kausalanalyse

Paneldaten gelten als zweitbeste Lösung für die Klärung von kausalen Zusammenhängen. Die beste wären kontrollierte Experimente. Der Grund ist, dass in Paneldaten zeitliche Abläufe bei der einzelnen Stichprobeneinheit nachvollzogen werden können, und das in Kombination mit der oftmals plausiblen Annahme, dass vieles an ihr nicht eigens beobachtet werden müsse, weil es sich im Zeitverlauf ohnehin nicht ändere und daher keine Auswirkung auf das haben könne, was sich im Zeitverlauf ändert. Paneldaten eignen sich somit vor allem, wenn erkennbar werden soll, welche Folgen eine Zustandsänderung bei einer Person zeitigt, etwa ein Übertritt aus Ausbildung in Beschäftigung, aus Beziehung in Ehe oder ein Umzug von einem Ort an einen anderen. Man hat dann bei der Erklärung der Folgen die Gewissheit, dass sie nicht dadurch beeinträchtigt werden, dass der eine Zustand bei einer Person und der andere bei einer anderen beobachtet wurde und sich diese beiden Personen in Dingen unterscheiden, welche die Folgen ebenfalls beeinflussen. So könnte die eine schulisches Lernen eine ganz tolle Sache finden, die andere es aber verabscheuen, sodass Ausbildung die eine und Beschäftigung die andere glücklicher macht. Vergleiche man die beiden Personen, während jene mit Lernneigung in Ausbildung und jene mit Lernabneigung in Beschäftigung ist, würde man womöglich keinen Unterschied im Ausmaß des Glücks feststellen, aber wenn man beiden beim Wechsel von Ausbildung in Beschäftigung zusähe, wie Paneldaten das ermöglichen, dann könnten sich bei beiden erhebliche Veränderungen des Glücks zeigen.

Das Beispiel mit dem Umzug zeigt aber auch, dass die Folgen nicht unbedingt aus einer Zustandsänderung der Person erwachsen müssen. Sie ist vorher so beschäftigt und so verheiratet wie nachher, begibt sich durch den Umzug aber aus einem Satz an äußeren Umständen in einen

anderen. Es ist die Veränderung in den äußeren Umständen, die Folgen zeitigt, keine Merkmalsveränderung bei der Person. Dennoch ist das selbstverständlich als Zustandsänderung der Person darstellbar. Geändert haben sich die Angaben über das Wohnumfeld der Person, die genauso an der Person haften wie der Beschäftigungsstatus oder der Familienstand. (Geändert haben sich wahrscheinlich auch die Angaben über die Wohnsituation selbst, und die Person hat jetzt zudem das Merkmal, umgezogen zu sein, das sie zuvor nicht hatte, alles Einflüsse, die befristet oder auf Dauer ihr Wohlbefinden steigern oder verringern können).

Giesselmann & Windzio (2014:97) weisen darauf hin, dass sozialwissenschaftliche Hypothesen zumeist die Effektwirkung einer Variable thematisieren und daher in einer „längsschnittlichen Semantik angelegt sind: Sie treffen Aussagen darüber, wie sich die abhängige Variable verändert, wenn sich die unabhängige Variable von einem Ausgangswert entfernt.“ Solche Aussagen können aus der statistischen Analyse von Querschnittsdaten jedoch niemals gewonnen, sondern ihnen immer nur nachträglich interpretativ aufgepfropft werden. Solange dies in vollem Bewusstsein der rein hypothetischen Natur solcher Aussagen geschieht, ist es sicherlich nicht nur zulässig, sondern auch produktiv, aber es wird zum unentschuldbaren Fehler, sobald dieses Bewusstsein fehlt oder aussetzt. Dagegen wird in einer Analyse von Veränderungen der Merkmale einer Stichprobeneinheit im Zeitverlauf „diese Semantik ..., zumindest näherungsweise, aufgegriffen“ (Giesselmann/Windzio 2014:97).

Paneldaten bringen nicht viel, wenn man, zum Beispiel, wissen will, wie sich eine optimistischere im Vergleich zu einer weniger optimistischen Lebenshaltung auf die Einkommensverläufe von Personen auswirken, solange die Lebenshaltungen sich nicht ändern. Der Vorteil von Paneldaten für die Analyse besteht einzig und allein darin, das Vorher und das Nachher einer Veränderung bei ein und derselben Stichprobeneinheit beobachten zu können und so ein hohes Maß an Gewissheit über die Ursache all dessen zu gewinnen, was im Nachher anders ist als im Vorher.

Von dem zentralen analytischen Vorteil abgesehen haben Paneldaten auch noch Vorteile betreffend die Validität von zeitlich auseinander liegenden Angaben zur selben Stichprobeneinheit. Eine häufige diesbezügliche Verwendung besteht in der Kombination von auf ein und dieselbe Person bezogenen Daten aus verschiedenen Zeitpunkten in einer einzigen Querschnittsuntersuchung, etwa wenn Zustände zu einem früheren Zeitpunkt als einflussreich auf aktuelle Zustände verstanden werden. Angaben, die eine Person über die Vergangenheit macht, sind mit einem erhöhten Risiko behaftet, dass sie falsch sein könnten. Eine Angabe über die beruflichen Ziele im Alter von 20 Jahren wird zutreffender sein, wenn sie von der Person erfragt wurde, als sie 20 war, als wenn sie erst mit 40 danach gefragt wurde. Hat man die Person zu mehreren Zeitpunkten befragt, so

wird es unnötig, sie über Zustände und Ereignisse der Vergangenheit zu befragen. Dadurch können Paneldaten auch für Querschnittsanalysen nützlich sein (Giesselmann/Windzio 2014:101).

Die kausalanalytische Absicht und die zutreffenderen Angaben über Zustände in der Vergangenheit haben in Deutschland dazu geführt, dass in den 2000er Jahren neben dem schon lange etablierten Sozio-ökonomischen Panel mehrere neue Panelprojekte auf den Weg gebracht worden, so das Nationale Bildungspanel, das Deutsche Familien- und Beziehungspanel und das Panel Arbeitsmarkt und soziale Sicherung (Brüderl 2010:963f; Giesselmann/Windzio 2014:96).

2.2. Stichprobeneinheit

Die Einheit, deren Merkmale im Zeitablauf beobachtet werden, muss nicht unbedingt die Person sein. Dazu eignen sich grundsätzlich auch Personengruppen, etwa der Haushalt, wenn man Vorkehrungen trifft, die dafür sorgen, dass Zugänge und Abgänge im Haushalt und altersbedingte Statusänderungen innerhalb des Haushalts keine Auswirkungen auf die abhängige Variable haben können. Würde man die Bildungsbeteiligung im Haushalt modellieren wollen, so würde sie sich einfach dadurch Erhöhen, dass ein Haushaltsmitglied in das Pflichtschulalter kommt, oder der Eintritt eines weiteren Haushaltsmitglieds in das Erwerbsalter könnte unmittelbar die Erwerbsbeteiligung bzw. das Haushaltseinkommen steigern. Man wird daher nicht den ganzen Haushalt, sondern die von Alter und Status her relevanten Teile als Einheit heranziehen und das Vorhandensein weiterer Haushaltsmitglieder eventuell als Einfluss auf das Verhalten der Einheit in die zu prüfende Hypothese einbeziehen.

Paneldaten spielen unter anderem eine bedeutende Rolle in Staatenvergleichen. Dabei sind die Staaten die Individuen, an denen man jene Merkmals- und Verhaltensänderungen im Zeiterlauf beobachtet, von denen man erwartet, dass sie einander beeinflussen. Es kommt nicht von ungefähr, dass man hierbei von Staaten spricht statt von Gesellschaften. Den Staat stellt man sich dabei nicht als Bevölkerung dar, sondern als Regierung und Gesetzgebung mit Eigenschaften, die jenen eines unsterblichen Individuums ähneln. Mit Bevölkerungen oder Bevölkerungsteilen müsste man dagegen auf die gleiche Weise verfahren wie mit Haushalten und sich überlegen, welche Teile davon für jenes Verhalten, das zu modellieren ist, überhaupt in Frage kommen.

Will man den Zeitverlauf des Verhaltens von Bevölkerungsteilen innerhalb einer Gesellschaft beobachten und erklären, dann muss man diese Bevölkerungsteile definieren und als Einheiten heranziehen. Ein Vergleich zwischen Einheimischen und Einwanderern stößt dann auf das Problem, dass dies nur zwei Einheiten sind, also eine Fallzahl von 2 vorliegt, während man vielleicht zehn

oder zwanzig Merkmale von jedem der beiden Bevölkerungsteile beobachten will. Das scheitert. Ein Mindestanforderung jeder statistischen Analyse ist, dass die Fallzahl größer ist als die Zahl der Merkmale. Die Vorgangsweise wird daher eine andere sein müssen. Man wird stattdessen Personen oder Haushalte als Einheiten definieren und ihre staatliche Zugehörigkeit als erklärende Variable in der Analyse berücksichtigen. Damit ändert sich auch das operative Ziel der Analyse. Falls es systematische Unterschiede zwischen Einheiten mit der einen und jenen mit der anderen staatlichen Zugehörigkeit gibt, muss sie nun darauf ausgerichtet sein, jene Merkmale zu finden, welche zu diesen Unterschieden führen. Erreicht ist dieses Ziel dann, wenn die Einflussgröße „staatliche Zugehörigkeit“ keinen nachweisbaren Einfluss mehr auf das Zustandekommen der Unterschiede hat, sondern die Unterschiede gänzlich auf andere Merkmale zurückgeführt werden können. Das Wort „gänzlich“ im voranstehenden Satz hat seine Berechtigung darin, dass die staatliche Zugehörigkeit ein rein nominelles Merkmal ist. Sie hat, so wie „weiblich“ oder „männlich“, per se keinen Inhalt, sondern trifft nur eine Unterscheidung. Als solche kann sie nichts erklären und beeinflussen. Einen Inhalt kann man in sie nur interpretierend hinein legen. Bei der staatlichen Zugehörigkeit ist der hinein interpretierte Inhalt oftmals „Kultur“. Das ist aber rein spekulativ. Es könnte auch etwas anderes sein. Wollte man den realen Einfluss von „Kultur“ in Erfahrung bringen, dann müsste man zuerst definieren, was darunter zu verstehen sei, als nächstes müsste es gemessen werden, und schließlich müssten die Messwerte in der statistischen Analyse auf ihren Erklärungswert hin getestet werden. Der zuvor der staatlichen Zugehörigkeit zugeordnete Einfluss müsste sich nun verringert haben oder ganz verschwunden sein. Wenn entlang des Geschlechts, des Alters oder diverser anderer Äußerlichkeiten Unterschiede auftreten, dann muss man ebenso verfahren: welche sachlichen Unterschiede verbinden sich mit dem Geschlecht, z.B. durchschnittliche Körpergröße, daher die Körpergröße aller in Erfahrung bringen, neben dem Geschlecht die Körpergröße in der Analyse berücksichtigen und sehen, ob nun noch ein Einfluss des Geschlechts verbleibt. Wenn ja, dann war die Körpergröße offenbar nicht alles, was in dem Hüllbegriff „Geschlecht“ drinsteckt, und man braucht weitere konkrete und somit letztlich messbare Größen, die mit Geschlecht in Verbindung stehen (könnten). Insofern diese sich als erklärungsstark herausstellen, wird man sich fragen, ob sie das berechtigterweise sind. Warum, etwa, sollte jemand aufgrund von 5cm oder 10cm zusätzlicher Körpergröße merklich besser bezahlt werden, wenn es im Job auf die Körpergröße nicht ankommt? Auch bei dem unerklärten Rest an Erklärungskraft, der am Geschlecht oder eben der staatlichen Herkunft hängen bleibt, wird man sich fragen müssen, welche Berechtigung er haben kann, aber mehr noch, worauf er beruht und wie er zustande kommt.

2.3. Fixed Effects

Der in der Methodenliteratur am breitesten besprochene Ansatz zur kausalanalytischen Nutzung von Paneldaten kreist um das Stichwort „fixed effects regression“. Die formale Herleitung der Fixed Effects Regression wird an dieser Stelle übersprungen. Sie kann in zahlreichen Lehrbüchern und Artikeln nachgesehen werden (unter anderem bei Brüderl 2010; Giesselmann/Windzio 2014; Giesselmann/Windzio 2012; Halaby 2004; Wooldridge 2013, 2010). Das Prinzip ist, dass eine Regressionsgleichung für die Durchschnittswerte einer Person über alle Zeitpunkte von der analogen Regressionsgleichung für einen bestimmten Zeitpunkt subtrahiert wird. Alle zeitkonstanten unabhängigen Variablen kürzen sich dabei heraus, wobei der entscheidende Punkt ist, dass sie das tun, ganz gleich ob sie beobachtet wurden oder nicht. Alle über die Zeit unveränderlichen Einflüsse sind damit aus der Analyse eliminiert. Das Verfahren ist in seiner Durchführbarkeit somit davon abhängig, dass nicht nur die abhängige Variable sich im Zeitverlauf ändert, sondern dass auch unter den unabhängigen zeitveränderliche sind, eine Einschränkung, der Querschnittsanalysen nicht unterliegen.

Die Auswirkungen von nicht beobachteten, zeitveränderlichen Variablen können mit der Subtraktionsmethode jedoch nicht bereinigt werden. Sie bleiben im Fehlerterm der Regression und verzerren die geschätzten Koeffizienten jener unabhängigen Variablen, mit deren Werten sie korreliert sind.

Eine Sonderform davon ist das „first-difference model“. Dabei wird nicht ein Regressionsmodell für die Durchschnittswerte, sondern für den ersten Messzeitpunkt vom selben für einen späteren Messzeitpunkt subtrahiert (Brüderl 2010:974; Wooldridge 2013).

Ein wiederholt auftretender Fehler, auch in begutachteten Artikeln, ist, dass trotz Anwendung einer Fixed Effects Regression zeitkonstante Variablen als Kontrollvariablen eingebaut werden. „Gerade die perfekte Kontrolle zeitkonstanter Merkmale wird aber im Rahmen einer within-Regression geleistet, der Ausfall zeitkonstanter Variablen stellt somit keinen Verlust dar, sondern indiziert vielmehr das zentrale, auch explizit motivierte Potenzial von Paneldaten zur Kausalitätsannäherung. Zudem lassen sich im FE Modell Interaktionen zeitkonstanter und zeitveränderlicher Variablen problemlos spezifizieren (und so die Mechanismen aufklären, die zu Gruppenunterschieden führen können)“ (Giesselmann/Windzio 2014:107, 97; siehe Brüderl 2010:975f; Halaby 2004:523ff). Mit anderen Worten, will man wissen, ob und wie sich das zeitkonstante Merkmal Geschlecht auf das Einkommen auswirkt, kann man es nicht direkt in die Regression aufnehmen, sondern muss es mit einem zeitveränderlichen Merkmal unter den unabhängigen Variablen multiplizieren. Sichtbar wird

dann, wie sich das Geschlecht auf die Einkommenswirksamkeit dieses einen Merkmals auswirkt bzw., wie man sagt, seine Wirksamkeit „moderiert“. „Zweitens ermöglichen sogenannte hybride Regressionen die gemeinsame Spezifikation von within- und between-Effekt. Solche Modelle erlauben also die Modellierung zeitkonstanter Variablen bei gleichzeitiger Spezifikation des within-Schätzers zeitveränderlicher Variablen. Besteht demnach im spezifischen Analysefall nicht nur ein Interesse an den Koeffizienten zeitveränderlicher Variablen, sondern ein zusätzliches substanzielles Interesse an den Effekten oder Mediationseigenschaften zeitkonstanter Merkmale, so wäre die Verwendung der Hybrid-Regression angezeigt“ (Giesselmann/Windzio 2014:107; Beispiele bei Brüderl 2010:976f und Halaby 2004:350ff).

2.4. Mehrebenenanalyse

Vor allem in den Biowissenschaften und der Soziologie wird die Grundidee der Fixed Effects Analyse häufig in Form sogenannter Mehrebenenanalysen umgesetzt (Hosoya u.a. 2014; Field 2009; Twisk 2007). Als Teil des weiten Feldes der Gemischten Modelle sind sie sehr flexibel im Einsatz, verlangen aber auch einige Genauigkeit. Zudem führen terminologische Mehrgleichigkeiten, in denen ähnlich klingende Ausdrücke unterschiedliche Bedeutungen haben, rasch zu Missverständnissen. Der Grundgedanke ist hier nicht einer der Subtraktion, sondern der Addition. Wenn, etwa, Personen über Zeit beobachtet werden, dann wird zunächst eine Regressionsgleichung für die zu allen Zeitpunkten gemessenen Werte formuliert, wie man das in einer Querschnittsanalyse ebenso täte. Dann kommt aber der Gedanke hinzu, dass die dabei ermittelten Koeffizienten beeinflusst sein können durch Merkmale der jeweiligen Person, zu der die Messungen auf der Zeitebene gehören. Daher wird auf der Personenebene eine zweite Regressionsgleichung formuliert, in der ein Koeffizient von der Zeitebene die abhängige Variable ist, und in die erste eingefügt. Für jeden auf der Zeitebene auftretenden Koeffizienten einschließlich der Regressionskonstanten wird auf der Personenebene eine eigene Regressionsgleichung formuliert, und sie alle werden in die zuvor auf der Zeitebene formulierte Gleichung eingefügt (Hosoya u.a. 2014; Langer 2009).

- Wenn nur die Gleichung für die Regressionskonstante eingefügt wird, so spricht man von einem Random Intercept Modell;
- Wenn nur die Gleichungen für die Steigungskoeffizienten eingefügt werden, so spricht man von einem Random Slope Modell;
- Wenn beides, dann ist die Rede von einem Random Intercept Random Slope Modell;
- Kommt die Zeit explizit und mit Random Slope vor, dann spricht man von einem Wachstumskurvenmodell.

Die Analogie zum Fixed Effects Verfahren besteht bei Random Slope Modellen: „Ein weiteres [der FE Analyse] äquivalentes Verfahren ist eine Regression mit personenspezifischen Koeffizienten. Dabei berechnet man im Prinzip mit den Beobachtungen einer jeden Person eine eigene Regression. Das gewichtete Mittel der Steigungskoeffizienten dieser Regressionen ist dann der FE-Schätzer. Für Personen, die während ihrer Beobachtungsdauer keine Varianz auf einer unabhängigen Variablen aufweisen, kann keine Regressionsgerade geschätzt werden, weshalb sie nicht in die Berechnung eingehen. ... macht dieser Schätzer noch einmal klar, dass es sich um eine reine Within-Betrachtung handelt: Es wird gefragt, wie sich Y ändert, wenn sich X ändert und zwar bei einer Person über die Zeit. Das Mittel der Antworten auf diese Frage ist dann der FE-Schätzer“ (Brüderl 2010:973f). Reine Random Intercept Modelle dagegen haben keine Fixed Effects Eigenschaften (Brüderl 2010:972).

Die soeben erwähnten Wachstumskurvenmodelle sind in einem Fall, wie dem vorliegenden, bei dem es mit dem Einkommen um eine abhängige Variable geht, von der im Zeitverlauf dem Trend nach Zunahmen zu erwarten sind, von besonderem Interesse. Ihr Charakteristikum ist die Einbeziehung der Zeit als unabhängige Variable (Hosoya u.a. 2014:204). Unter Einsatz eines solchen Modells lässt sich prüfen, ob die Haushalte signifikant voneinander abweichende Zeitverläufe des Einkommens aufweisen.

Die Erkenntnis, dass unter mehr als 2,5 Millionen Haushalten mit mindestens einer Person in erwerbsfähigem Alter signifikante Unterschiede der Einkommensentwicklung anzutreffen sind, würde wenig überraschen. Zudem ist die Forschungsfrage ja spezifischer, nämlich ob die Einkommen von Kategorien von Haushalten, die an der staatlichen Herkunft gebildet sind, sich im Zeitverlauf aufeinander zu oder voneinander weg bewegen. Man muss also in der Regression die Kategoriezugehörigkeit als Variable berücksichtigen. Die staatliche Herkunft eines Haushalts ist aber quasi-zeitkonstant und hat daher in einem Fixed Effects Modell keine Daseinsberechtigung. Ein unüberwindliches Hindernis ist dies zum Glück nicht. In Mehrebenenanalysen lassen sich, wie schon erwähnt, zeitkonstante Merkmale ohne allzu große Herausforderung in eine Hypothese einbauen. Das geschieht, indem sie mit einer zeitveränderlichen Variable bzw. in einem Wachstumskurvenmodell mit der Zeitvariable selbst multiplikativ verbunden („interagiert“) werden. Man muss dann die zeitveränderliche Variable – etwa die Zeitangabe selbst – und die zeitkonstante Variable zusammen mit ihrer Interaktion in die Regressionsgleichung aufnehmen (Hosoya u.a. 2014:204f). (Sind die Variablen kategorial, braucht man nicht alle drei, sondern nur die zeitveränderliche und die Interaktion). Mit steigender Zahl an Variablen steigt aber das Risiko, dass ein Mehrebenenmodell nicht mehr konvergiert, dass also kein (zutreffendes) Ergebnis ermittelt werden kann. Das

kann ein Hinweis sein, dass die Regressionsgleichung überarbeitet werden müsste (Pötschke 2014). Wenn dazu kein Anlass besteht, dann kann eine Lösung sein, die staatliche Herkunft nicht interagierend in die Regressionsgleichung einzubauen, sondern die Regression für jede staatliche Herkunft separat durchzuführen. Dieser Weg wird im vorliegenden Fall gewählt.

3. Stichprobe

Die Daten von SILC 2008 bis 2013 enthalten inklusive der Wiederholungen 26.937 Haushalte mit mindestens einer Person in erwerbsfähigem Alter.

Für die Längsschnittanalysen wurden nur jene Haushalte mit mindestens einer Person in erwerbsfähigem Alter herangezogen, die in den Jahren 2008, 2009 und 2010 neu in die Stichprobe kamen. War niemand mehr in erwerbsfähigem Alter, wurde der Haushalt in der Auswertung nicht mehr berücksichtigt, auch wenn er noch an weiteren Folgebefragungen teilgenommen hatte.

3.1. Stichprobenschwund

In der vierten Befragungswelle sind nur mehr etwa 63% der Haushalte weiterhin in der Stichprobe enthalten, in der dritten Welle sind es noch 71%, in der zweiten noch etwa 80%. Zum Teil geht der Schwund darauf zurück, dass Haushalte während der Panellaufzeit aus dem Erwerbsalter hinaus altern.

Tab. 1: Anzahl der Haushalte in der Stichprobe

Welle	Anzahl			Prozent			Gesamt
	Erstbefragungsjahr			Erstbefragungsjahr			
	2008	2009	2010	2008	2009	2010	
1	1.937	2.084	2.064	100,0	100,0	100,0	100,0
2	1.473	1.725	1.694	76,0	82,8	82,1	80,4
3	1.280	1.531	1.508	66,1	73,5	73,1	71,0
4	1.115	1.370	1.326	57,6	65,7	64,2	62,6

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

3.2. Bestimmungsgründe des Schwundes

Eine lineare Regression der Anzahl der Befragungswellen, an denen ein Haushalt beteiligt war, auf eine Anzahl unabhängiger Variablen ergibt ein ambivalentes Ergebnis. Zwar wird eine Reihe von Variablen auf 95%-Niveau signifikant, darunter das Äquivalisierte Einkommen, das Alter, die Bildung, alle drei samt Quadrat (stets konkav), ausländische Haushalte (negativ), Kleingemeinden, das Erstbefragungsjahr, die Erwerbsbeteiligung (negativ) und die Zufriedenheit mit dem Leben, aber sie tragen insgesamt so gut wie nichts dazu bei, wie lange ein Haushalt in der Stichprobe ausharrt. Im ungewichteten Fall erzielt die Regression ein korrigiertes R-Quadrat von lediglich 0,048, im gewichteten von 0,043. Das Ergebnis von entsprechenden logistischen Regressionen blieb ebenso dürftig. Es wurde daher auf eine Korrektur der Verzerrung, die durch teilweise Selbstauswahl in die Stichprobe zu befürchten wäre, verzichtet.

4. Deskriptive Ergebnisse

4.1. Das äquivalisierte Nettohaushaltseinkommen pro Jahr

4.1.1. Die Variable

Das äquivalisierte Nettohaushaltseinkommen liegt in den Daten bereits in berechneter Form vor. Berechnet wird es, indem das Nettohaushaltseinkommen durch die bedarfsgewichtete Haushaltsgröße dividiert wird (siehe unten). Leider entsteht dabei ein Messfehler, weil das Einkommen jenes des Vorjahrs ist, die Haushaltsgröße aber jene des Erhebungszeitpunkts. Für die Befragungswellen 1 bis 3 ließe sich das korrigieren, weil bei ihnen aus den jeweils folgenden Wellen 2 bis 4 das möglicherweise etwas besser zur Haushaltsgröße passende Haushaltseinkommen bekannt ist, sodass das Haushaltseinkommen durch ein Bedarfsgewicht aus dem gleichen Jahr dividiert werden könnte, aber für die letzte Befragungswelle ist das nicht möglich, sodass es besser ist, keine Neuberechnung vorzunehmen.

In einer anderen Hinsicht wurden aber Neuberechnungen vorgenommen. Es kommt vor, dass Personen im Panelverlauf den Haushalt wechseln, etwa weil sie sich aus dem bisherigen herauslösen und einen eigenen gründen. Für die vorliegende Analyse wurden solche Trennungen rückgängig gemacht und die benötigten Haushaltsvariablen neu berechnet.

4.1.2. Staat

Jede der drei Variablen „Geburtsstaat“, „Staatsangehörigkeit“ und „frühere Staatsangehörigkeit“ steht in SILC in sechs Ausprägungen zur Verfügung:

1. Österreich
2. Übrige EU15 und EFTA
3. Alle weiteren EU Mitgliedsstaaten zum Erhebungszeitpunkt
4. Ehemaliges Jugoslawien ohne Slowenien
5. Türkei
6. Sonstige Staaten.

Die Variable „Staat“ wurde gebildet, indem der höchste in den drei Variablen auftretende Wert herangezogen wurde. Sie wurde jedes Jahr für jede Person im Haushalt einzeln berechnet. Das heißt im Extremfall, dass eine in Österreich aufgewachsene Person mit österreichischer Staatsangehörigkeit, die bis zur Volljährigkeit eine andere Staatsangehörigkeit hatte, hier mit ihrer früheren Staatsangehörigkeit dabei ist. Es gab zwei Gründe, so vorzugehen. Der inhaltliche ist, dass Mikrozensusauswertungen immer wieder zeigen, dass der Herkunftsstaat der Eltern für den weiteren Erfolg im österreichischen Beschäftigungswesen relevant bleibt. Das heißt, die Auswertung wurde nicht als reines Einwandererpanel, sondern mehr als Integrationspanel angelegt. Der mehr technische Grund ist, dass volle Konzentration auf Geburt im Ausland, wie ein striktes Einwandererpanel das tun müsste, die Fallzahlen merklich verkleinert hätte.

Eine vierte Variable, die ebenfalls in denselben sechs Ausprägungen zur Verfügung steht, ist die „zweite Staatsangehörigkeit.“ Sie wurde jedoch nicht in die Berechnung der Variable „Staat“ einbezogen.

Für die Analyse herangezogen wurde der Staat jener Person, die im Haushalt als Hauptverdiener*in designiert war, und zwar in nur drei Ausprägungen:

- Österreich
- EU15/EFTA
- Sonstige.

Wie alle Variablen in der Untersuchung, wurde auch der Staat für jedes Erhebungsjahr, an dem ein Haushalt beteiligt war, neu berechnet. Dadurch kann es vorkommen, dass er sich in einem Haus-

halt im Lauf der vier Befragungen ändert. Giesselmann & Windzio (2014:102) unterscheiden bei den unabhängigen Variablen in Panelstudien zwischen zeitkonstanten, quasi-zeitkonstanten, zeitveränderlichen und Trendvariablen, wobei die letzten ein Sonderfall der zeitveränderlichen sind. „Als quasi-zeitkonstante Variablen gelten solche Merkmale, die zwar konzeptionell als zeitinvariante Merkmale angelegt sind, aber trotzdem marginale Variation auf individueller Ebene aufweisen“ (Giesselmann/Windzio 2014:103). Auf sie ein Fixed Effects Verfahren anzuwenden kann irreführende Schätzgrößen hervorbringen (Giesselmann/Windzio 2014:101). Der Staat in seinen drei Ausprägungen ist eine solche quasi-zeitkonstante Variable. Unter den 13.974 in die Analyse einbezogenen Haushalt-Jahr-Einheiten gibt es 9.477 der Wellen 2 bis 4 und darunter 172 (1,8%), in denen ein Haushalt eine andere Staatskennung hat als im Jahr davor. Die Wechsel entstehen durch Wechsel der Hauptverdienerrolle im Haushalt bzw. eventuell auch durch bei derselben Person zwischen den Jahren differierende Angaben über den Geburtsstaat, die Staatsangehörigkeit und die frühere Staatsangehörigkeit.

4.1.3. Der Befund

Zwischen den SILC-Runden 2008 und 2013, also zwischen den Einkommensjahren 2007 und 2012, nahm das durchschnittliche äquivalisierte Haushaltsnettoeinkommen in Haushalten mit Hauptverdienerin bzw. Hauptverdiener aus Österreich um trendgemäß €584 pro Jahr zu. Das entspricht 2,3% ihres Durchschnittseinkommens über den ganzen Zeitraum (Tab. 2). Im Vergleich dazu stieg das äquivalisierte Haushaltsnettoeinkommen jener Haushalte mit Hauptverdienerin bzw. Hauptverdiener aus Österreich, die zwischen 2008 und 2010 neu in die SILC-Stichprobe kamen und in den drei Folgejahren möglichst vollständig wieder befragt wurden, um durchschnittlich €815 oder 3,1% ihres Durchschnittseinkommens während der Periode (Tab. 3, Abb. 1). Zieht man nur jene Haushalte heran, die bei allen vier Befragungswellen befragt werden konnten – und das sind genau jene, die in der 4. Welle noch vorhanden sind, weshalb in Tabelle 3 in der entsprechenden Zeile die Werte im linken Teil und im rechten Teil übereinstimmen müssen, dann ist deren durchschnittliches äquivalisiertes Haushaltsnettoeinkommen bei der Erstbefragung um etwa €300 pro Jahr höher als unter Einbeziehung auch jener, die nur zwei- oder dreimal befragt werden konnten, sodass die Trendsteigerung nur €722 pro Jahr beträgt bzw. 2,8% (Tab. 3). Auf die beiden anderen Herkunftsgebiete und die Zuzugsperioden kommt die Sprache weiter unten.

Tab. 2: Äquivalisiertes jährliches Nettohaushaltseinkommen

Jahr	Herkunftsstaat			Zuzug nicht EU15/EFTA		
	AT	EU15/EFTA	Sonst	1960-1985	1986-1995	1996-2005
2008	23.823	22.017	15.894			
2009	24.915	25.856	17.456	18.737	18.278	13.725
2010	25.810	24.840	17.803	18.926	19.425	12.820
2011	26.131	30.210	16.914	18.379	16.525	14.158
2012	26.516	28.377	17.780	20.690	17.201	14.819
2013	26.885	26.377	18.433	20.882	18.640	14.446
Steigung	584	992	365	605	-150	344
%	2,3	3,8	2,1	3,1	-0,8	2,5
n pro Jahr	Ca. 3500	130-165	525-560	57-84	159-178	80-89

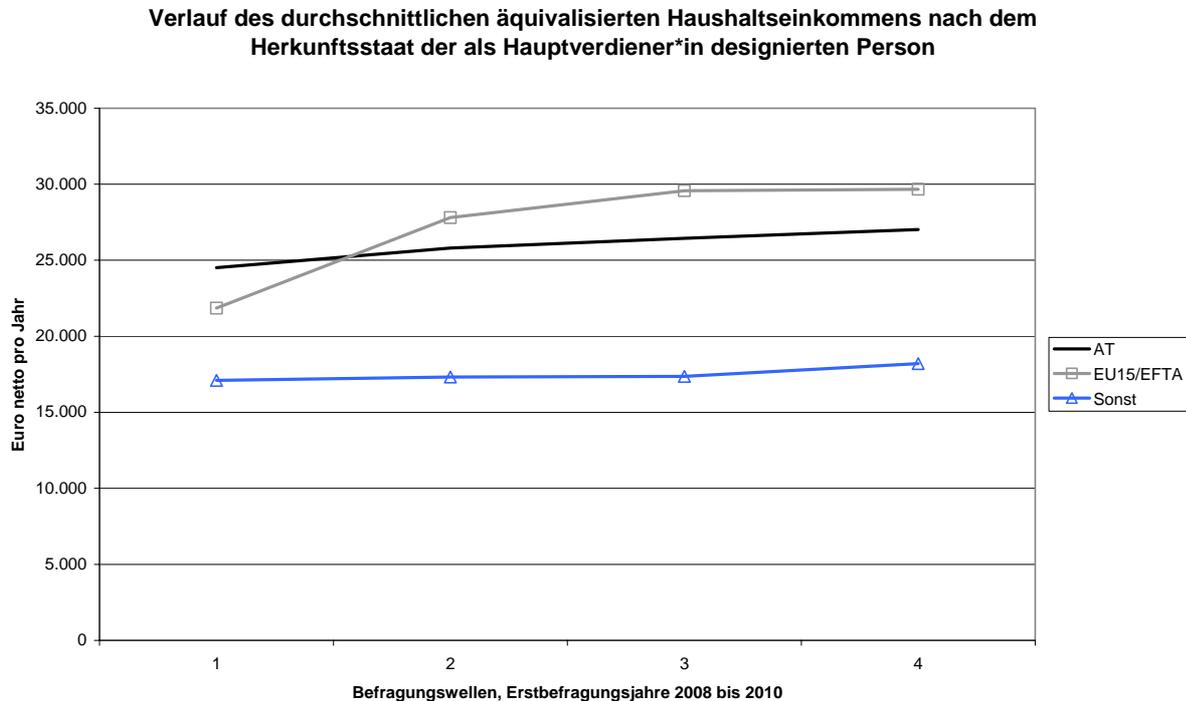
Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Tab. 3: Äquivalisiertes jährliches Nettohaushaltseinkommen

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	24.515	21.853	17.091	24.838	22.695	17.418
2	25.795	27.805	17.315	25.994	28.123	17.680
3	26.436	29.567	17.354	26.677	29.612	17.969
4	27.018	29.665	18.197	27.018	29.665	18.197
Steigung	815	2.520	336	722	2.240	263
%	3,1	9,3	1,9	2,8	8,1	1,5

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Abb. 1



Was ergibt sich daraus an Erkenntnissen über den Verlauf der äquivalisierten Nettojahreseinkommen der Haushalte mit Hauptverdienerin bzw. Hauptverdiener aus Österreich?

- Die Einkommenssteigerungen erweisen sich bei Panelbetrachtung als etwas größer als bei Betrachtung des bloßen Zeitverlaufs.
- In der Panelbetrachtung sind die Unterschiede zwischen den Wachstumsraten bedeutend größer als im einfachen Zeitverlauf.
- Die Haushalte, die an allen vier Wellen teilnahmen, sind offenbar nur geringfügig einkommensstärker als jene, die im Lauf der Zeit unauffindbar wurden, sich nicht mehr befragen ließen, oder niemanden mehr in erwerbsfähigem Alter enthielten. Der stärkere Anstieg (3,1% statt 2,8%), wenn alle in der Panelstichprobe enthaltenen Haushalte in die Auswertung einbezogen werden, schuldet sich dem sukzessiven Ausstieg der weniger einkommensdynamischen Haushalte, also vermutlich vor allem solchen, die sich der Pension nähern oder vor Ende des erwerbsfähigen Alters in Pension gehen.

Die EU15/EFTA Spalten der beiden Tabellen sind der Vollständigkeit halber angeführt, aber die Zahl der Fälle ist relativ gering – in der ersten Panelwelle ergeben die drei Stichprobenkohorten zusammen nur 142 Haushalte und die Zahl sinkt bis zur vierten Welle auf 95, weshalb es teils zu

sprunghaften Zu- und Abnahmen des äquivalisierten Haushaltseinkommens kommt (Tab. 2), die vor allem als statistische Schwankungen zu verstehen sind. In der Panelbetrachtung gibt es diese Schwankungen jedoch nicht (Tab. 2), was zeigt, dass nicht die einzelnen Haushalte ihnen unterliegen, sondern nur der Durchschnitt über alle aus EU15/EFTA Staaten stammenden Haushalte einschließlich jener, die vor 2008 und nach 2010 zuzogen. Was sich im Panel dagegen sehr deutlich abzeichnet ist die Verlangsamung der Zuwächse des äquivalisierten Haushaltseinkommens bis hin zum Stillstand (Tab. 3, Abb. 1).

Das jährliche äquivalisierte Nettoeinkommen der Haushalte mit Hauptverdienerin bzw. Hauptverdiener von außerhalb der EU15/EFTA Staaten stieg über die Jahre um durchschnittlich 2,1%, also etwas langsamer als bei den Haushalten aus Österreich (Tab. 2). Im Panel ist das deutlicher, denn dort sind es nur 1,9% im Vergleich zu 3,1% bei den Haushalten aus Österreich. Der durchschnittliche Einkommenszuwachs betrug nur €336 pro Jahr statt €815 trotz des viel niedrigeren Ausgangswerts von rund €17.100 statt €24.500. Die Lücke vergrößert sich von rund €7.500 in Welle 1 auf rund €9.200 in Welle 4.

4.1.4. Aufenthaltsdauer

Die Fallzahl lädt nicht unbedingt dazu ein, aber man wird sich sicherlich fragen, inwiefern die Aufenthaltsdauer der Haushalte eine Rolle für den Einkommensverlauf spielt. Die eine Antwort auf die Frage muss natürlich sein, dass das Panel gerade das anzeigt, nämlich was durch Verlängerung der Aufenthaltsdauer um drei Jahre mit dem Einkommen geschieht. Die Haushalte haben aber unterschiedliche der Panelkarriere vorangehende Aufenthaltsdauern. Das heißt, sie unterscheiden sich in einem zeitkonstanten Merkmal, nämlich dem Jahr des Aufenthaltsbeginns, bzw. in dem ebenso zeitkonstanten Merkmal, wie viele Jahre zwischen Aufenthaltsbeginn und Eintritt in die Panelstichprobe vergangen sind. Das Merkmal könnte einen Einfluss auf das Einkommen bei Eintritt in die Stichprobe haben, aber ebenso auch auf den Einkommensverlauf während der Panelzugehörigkeit. Beides kann aber auch durch das Alter bedingt sein. Für Regressionsanalysen, besonders auch für solche im Längsschnitt, wie hier, wirft dies Probleme auf. Faktum ist, dass Menschen vorwiegend im Alter zwischen 18 und 38 migrieren, sodass die Aufenthaltsdauer und das Alter miteinander korrelieren, im vorliegenden Fall mit etwa $r=0,7$. In der Regel muss man sich in Regressionsanalysen zwischen Aufenthaltsdauer und Alter entscheiden, weil die beiden so stark korreliert sind, dass Multikollinearität entsteht und dann nicht mehr mit Gewissheit gesagt werden kann, ob ein allfälliger Effekt auf das Einkommen von der einen oder von der anderen der beiden Variablen ausgeht oder von beiden gemeinsam. Bei der Aufenthaltsdauer müsste in einer Regressionsanalyse außerdem eine Entscheidung getroffen werden, wie bei all jenen, die seit Geburt im

Inland leben, mit der Aufenthaltsdauer verfahren werden sollte. Bei ihnen die Aufenthaltsdauer einfach mit dem Alter gleichzusetzen, würde die Korrelation zwischen Alter und Aufenthaltsdauer nochmals sehr stark erhöhen. Die eigentlich interessante Angabe, nämlich die Aufenthaltsdauer nicht der Person, sondern der Familie, die auch einige hundert Jahre betragen kann, ist nicht eruierbar. Man hat es hier mit vorerst unlösbaren Problemen zu tun.

Das Stichwort „Zeitkonstanz“ im vorigen Absatz verwies schon auf das weiter oben diskutierte Problem, dass Variablen, die sich über Zeit nicht ändern, für Panelanalysen genuiner Art keinen Mehrwert haben. Dieses Problem, immerhin, ist lösbar, und stellt somit nicht mehr als eine Herausforderung dar.

Man hat mit der Betrachtungsweise nach Zuzugskohorten ein Mittelding zwischen dem bloßen Zeitverlauf und der Panelbetrachtung vor sich. Es sind damit zumindest die Aufenthaltsdauer und die wirtschaftlichen, rechtlichen, sozialen und politischen Bedingungen zu Beginn des Aufenthalts einigermaßen fixiert und müssen nicht mehr eigens als Variablen bedacht werden. Mitunter kann damit auch eine gewisse soziale Einheitlichkeit der eingewanderten Bevölkerung einhergehen.

Der Zeitraum des Aufenthaltsbeginns wird in SILC seit 2009 erfragt. In den zur Analyse zugänglichen Daten ist angegeben, ob der Aufenthalt „vor 1960“ begann, „zwischen 1961 und 1970“ – das Jahr 1960 wurde vermutlich nicht vergessen, sondern die Bezeichnung des ersten Zeitraums müsste wahrscheinlich richtig „bis 1960“ lauten – zwischen 1971 und 1980, dann in Fünfjahresperioden bis 2010 und anschließend (vorerst) in Einzeljahren. Für unter 16-Jährige fehlt die Angabe. Da in einem Haushalt Personen mit unterschiedlichem Aufenthaltsbeginn zusammenleben können, muss definiert werden, wie der Aufenthaltsbeginn des Haushalts bestimmt werden soll. Im vorliegenden Fall wurde der früheste herangezogen, der sich unter den im Ausland geborenen Personen fand. Mehrpersonenhaushalte wurden nur dann als zugezogen gewertet, wenn mehr als eine Person im Haushalt im Ausland geboren war. Bei Einpersonenhaushalten stellte sich die Frage nicht. Dadurch verringert sich die Fallzahl auf die in den Tabellen 2, 5 und 7 für die drei zahlenmäßig wichtigsten Zuzugsperioden genannten Werte.

Hinsichtlich des Einkommens gab es je nach dem Zuzugszeitraum große Unterschiede, die allerdings wegen der geringen jährlichen Zahl an befragten Haushalten unsicher bleiben. Demnach nahmen die äquivalisierten Einkommen der Haushalte, deren Zuzug zwischen 1960 und 1985 stattfand, um durchschnittlich 7,6%, bei Zuzug um 1990 herum um 0,4% und bei Zuzug um 2000 herum um 4,6% (Tab. 2). Die zeitlichen Abgrenzungen sind hier nicht frei wählbar, weil die Kategorien in der Befragung vorgegeben sind, aber auch nicht ganz beliebig. Der Zeitraum 1986 bis 1995

entspricht sehr genau der Auflösung des Ostblocks und dem kriegerischen Zerfall Jugoslawiens, die gemeinsam die fünfte große Flüchtlingswelle seit 1945 erzeugten, und der Zeitraum 1996 bis 2005 ist genau jene Periode, in der die sechste große Flüchtlingswelle seit 1945 an- und abschwoll (fast genau 200.000 Asylanträge von 1998 bis 2005). Obwohl sie ziemlich sicher ähnliches Ausmaß hatten, ist die letztere in der EU-SILC Stichprobe seit 2009 mit jährlich nur rund 85 Haushalten nur halb so stark vertreten wie die frühere, die jeweils um die 170 Haushalte aufweist. Möglicherweise sind die um 1990 eingewanderten Haushalte dauerhafter in Österreich geblieben als die um 2000 zugeströmten, eine Frage, die sich grundsätzlich klären ließe.

4.1.5. Zuzugsperioden im Panel

Der Aufenthaltsbeginn wurde zwar erst ab 2009 erhoben, aber da er zeitkonstant ist, gilt die 2009 gemachte Angabe bei den Panelhaushalten auch für 2008. Nur für jene Haushalte, die 2008 in die Stichprobe kamen und danach kein zweites Mal mehr befragt werden konnten, liegt keine Angabe vor.

Um nun zur Hauptfrage zu kommen: Was sich nur beschränkt andeutet, ist ein Aufholen der Haushalte von außerhalb der EU15/EFTA Staaten gegenüber jenen mit Hauptverdienerin oder Hauptverdiener aus Österreich. Nur in der Zuzugskohortenbetrachtung wird – bei aller statistischen Unsicherheit – die Möglichkeit erkennbar, dass einzelne Zuzugskohorten ein rascheres Einkommenswachstum aufweisen könnten als die nicht eingewanderten Haushalte. Damit stellt sich quasi die Forderung in den Raum, auch die Panelbetrachtung nach Zuzugskohorten zu differenzieren. Von der Fallzahl her ist das selbstredend nicht ohne Risiko. In der vierten Befragungswelle stehen bei den bis 1985 zugezogenen noch 41 Haushalte zur Auswertung zur Verfügung, bei den 1986 bis 1995 zugezogenen immerhin noch 107, bei den 1996 bis 2005 zugezogenen aber nur 54. Ein Aufholprozess deutet sich hier bei den seit Mitte der 1990er Jahre zugezogenen Haushalten zwar an, aber die statistischen Unsicherheiten sind bei weitem zu groß, um definitive Aussagen treffen zu können. Den Anzeichen nach gilt: Während die nicht eingewanderten Haushalte über die vier Befragungswellen ein Einkommenswachstum von durchschnittlich rund €730 (2,9%) pro Jahr aufweisen, sinken die äquivalisierten Einkommen der bis 1985 zugezogenen Haushalte, jene der 1986 bis 1995 zugezogenen Haushalte steigen um rund €610 (3,3%), bleiben also zurück, und jene der 1996 bis 2005 zugezogenen Haushalte steigen um rund €1.600 (9,5%) pro Jahr. Wie gesagt, das ist alles mit großer Unsicherheit behaftet, weshalb auch keine Tabelle oder Abbildung geboten wird. Es gibt auch eine kleine Zahl von Haushalten – in der vierten Welle noch 15 – die zwischen 2006 und 2009 zugezogen sind. Ihr Einkommenswachstum beträgt im Durchschnitt rund €5.000 (19%) pro Jahr und ihr durchschnittliches Einkommen übertrifft ab der zweiten Welle dasje-

nige aller anderen Zuzugskohorten einschließlich der nicht migrierten Haushalte. Es steht zu bezweifeln, dass letzteres die Realität der seit 2006 erfolgten Einwanderung adäquat abbildet, wenngleich es eine Tatsache ist, dass sie ungewöhnlich häufig über Ausbildungen von der Matura aufwärts verfügt.

4.1.6. Periodeneffekte?

Für die Panelbeschreibung wurden drei Stichprobenkohorten zusammengefasst, nämlich die 2008, 2009 und 2010 neu in die SILC-Stichprobe eintretenden Haushalte. Das sind insofern keine gewöhnlichen Jahre, als 2009 eine tiefe wirtschaftliche Rezession zu verzeichnen war gefolgt von einer gewissen Erholung 2010. Die Frage ist, ob und wie sehr sich das in Unterschieden zwischen den drei Stichprobenkohorten niederschlägt. Der Erwartung nach wird es sich nicht niederschlagen, weil der Eintritt in die Stichprobe ja nicht durch besondere Ereignisse oder Merkmale der Haushalte ausgelöst wurde, sondern ein reines Zufallsprodukt war. Aus dem Stichprobeneintritt ergibt sich, beispielsweise, kein Hinweis auf das Zuzugsjahr. Der durchschnittliche Haushalt, der 2009 in die Stichprobe kam, sollte im Jahr 2009 nicht anders betroffen sein als ein ebenso durchschnittlicher, der 2008 in die Stichprobe kam. Beeinträchtigt wird die Beantwortung durch eben jenen Sachverhalt, der zur Zusammenfassung der drei Stichprobenkohorten Anlass gab, nämlich die relativ kleinen Fallzahlen der einzelnen Stichprobenkohorten. Das drückt sich in Tabelle 4 in den großen Schwankungsbreiten aus. Sie sind so groß, dass sich jede Aussage über Unterschiede zwischen den Stichprobenkohorten an sich verbietet. Der größte tritt bei den Haushalten aus den EU15/EFTA Staaten auf, wo die Einkommen der Kohorte 2009 weit über jenen der anderen beiden Kohorten liegen. Bei ihr ist die Fallzahl aber so klein – zwischen 26 und 48 befragte Haushalte – dass die Schwankungsbreite in den Wellen 2 bis 4 mehr als $\pm\text{€}10.000$ beträgt. Dazu tragen auch die größeren – aber nicht großen, denn der Variationskoeffizient ist maximal 0,33 – Standardabweichungen bei. Man kann nicht mit 95% Wahrscheinlichkeit davon ausgehen, dass die Einkommen der drei Kohorten tatsächlich voneinander verschieden sind.

Tab. 4: Verlauf des äquivalisierten Nettohaushaltseinkommens pro Jahr über die vier Panelwellen nach Jahr des Eintritts in die Stichprobe und Herkunftsstaat des Haushalts

Österreich	Stichprobenkohorte			plus/minus		
	2008	2009	2010	2008	2009	2010
1	23.525	24.610	25.379	794	720	860
2	25.556	25.949	25.860	1.082	879	823
3	25.988	26.546	26.713	956	914	960
4	26.334	27.220	27.395	1.067	1.031	1.047
EU15/EFTA						
1	20.810	25.148	19.841	3.502	7.269	4.907
2	23.533	31.016	27.332	3.935	10.389	3.976
3	25.026	39.777	22.322	4.413	11.307	3.640
4	26.939	39.655	22.998	4.561	12.554	5.522
Sonst						
1	16.218	16.252	19.034	1.293	1.538	2.428
2	17.162	17.383	17.396	2.218	1.885	1.264
3	15.323	18.287	18.284	1.402	1.682	1.663
4	16.166	18.896	19.333	1.420	1.566	1.873

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

4.2. Quellen des Wachstums des äquivalisierten Haushaltseinkommens

Das äquivalisierte Haushaltseinkommen entsteht aus der Division des Haushaltseinkommens durch die bedarfsgewichtete Haushaltsgröße. Zunahmen im Zähler ebenso wie Verringerungen im Nenner könnten die Steigerungen des äquivalisierten Haushaltseinkommens bewirkt haben. Wir betrachten zuerst das Einkommen, dann die Haushaltsgröße.

4.2.1. Das Nettohaushaltseinkommen eines Jahres

Seit 2015 stehen ab dem SILC-Erhebungsjahr 2008 Administrativdaten über die versteuerten Einkommen aus unselbständiger Erwerbstätigkeit und Befragungsdaten über die Entnahmen der Selbständigen aus dem Unternehmen zur Verfügung. Es handelt sich um das Einkommen des gesamten, dem Erhebungsjahr vorausgehenden Kalenderjahres.

Rein im Zeitverlauf stiegen die jährlichen Nettoeinkommen der Haushalte mit Hauptverdienerin bzw. Hauptverdiener aus Österreich um durchschnittlich 2,1% pro Jahr, mit Hauptverdienerin bzw. Hauptverdiener von außerhalb der EU15/EFTA Staaten um 1,9% (Tab. 5). Der Unterschied ist ähnlich wie beim äquivalisierten Haushaltseinkommen (Tab. 2). Das lässt erwarten, dass die Entwicklung der Haushaltsgrößen praktisch parallel sein wird.

In der Panelbetrachtung zeigt sich im Gegensatz dazu, dass das Nettohaushaltseinkommen der Haushalte aus Österreich im Durchschnitt über die vier Befragungswellen um 4,4% stieg, jenes der Haushalte von außerhalb der EU15/EFTA Staaten zwar um 4,9%, aber dennoch um einen um €300 pro Jahr geringeren Betrag (Tab. 6). Die Wachstumsraten sind geringer, wenn nur die Haushalte herangezogen werden, die in allen vier Wellen befragt werden konnten. Sie betragen dann 4,2% und 4,6%, wobei auch hier die größere Rate sich nicht aus höherem Wachstum, sondern aus dem viel niedrigeren Ausgangsniveau ergibt. Das Wachstum ist um rund €460 pro Jahr geringer (Tab. 6). Die Unterschiede der Wachstumsraten sind mit 0,5 bzw. 0,4 Prozentpunkten viel kleiner als beim äquivalisierten Haushaltseinkommen, wo sie 1,2 bzw. 1,3 Prozentpunkte betragen. Der größere Teil der Wachstumsunterschiede beim äquivalisierten Haushaltseinkommen der Panelhaushalte muss daher auf unterschiedliche Entwicklung der Haushaltsgrößen zurückgehen.

Zwischen den Zuzugskohorten sind die Unterschiede erheblich, aber statistisch schlecht abgesichert. Das Einkommenswachstum der Kohorte 1986-1995 scheint in den letzten Jahren geringer ausgefallen zu sein als jenes der beiden zeitlich benachbarten Kohorten, während ihr Einkommen die ganze Zeit über höher war (Tab. 5).

Tab. 5: Jährliches Nettohaushaltseinkommen

Jahr	Herkunftsstaat			Zuzug nicht EU15/EFTA		
	AT	EU15/EFTA	Sonst	1960-1985	1986-1995	1996-2005
2008	41.086	32.773	30.666			
2009	43.015	39.844	31.783	28.068	32.055	24.290
2010	44.356	37.599	31.903	28.798	32.678	22.943
2011	45.182	48.429	32.380	28.012	29.898	28.025
2012	45.755	44.946	33.525	30.520	31.618	27.302
2013	45.655	43.844	33.777	29.903	33.547	24.680
Steigung	911	2.328	607	539	192	514
%	2,1	5,6	1,9	1,9	0,6	2,0
n pro Jahr	Ca. 3500	130-165	525-560	57-84	159-178	80-89

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Tab. 6: Jährliches Nettohaushaltseinkommen

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	41.523	32.688	30.752	41.819	33.364	31.404
2	44.334	42.804	31.828	44.469	41.563	32.242
3	46.053	47.146	33.393	46.529	46.205	34.941
4	47.487	47.964	35.654	47.487	47.964	35.654
Steigung	1.961	5.017	1.627	1.906	4.844	1.545
%	4,4	11,8	4,9	4,2	11,5	4,6

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

4.2.2. Bedarfsgewichtete Haushaltsgröße

Die Bedarfsgewichtung erfolgt nach der von Eurostat verwendeten OECD-Skala (Hagenaars et al. 1994): eine Person ab 14 Jahren im Haushalt erhält das Gewicht 1, jede weitere Person ab 14 Jahren das Gewicht 0,5, jede Person unter 14 Jahren das Gewicht 0,3. Die bedarfsgewichtete Haushaltsgröße ist die Summe der Gewichte eines Haushalts.

Die Angabe bezieht sich auf den Erhebungszeitpunkt und ist daher nicht zeitgleich mit dem Haushaltseinkommen, das sich auf das vorangehende Kalenderjahr bezieht.

Die vorhin geäußerten Erwartungen bezüglich der Trends bei den Haushaltsgrößen bestätigen sich selbstverständlich. Im reinen Zeitverlauf zeigt sich kein Unterschied zwischen Haushalten aus Österreich und solchen von außerhalb der EU15/EFTA Staaten. Bei beiden schrumpft die Haushaltsgröße um durchschnittlich 0,2% pro Jahr (Tab. 7).

Ganz anders in der Panelbetrachtung. Hier nahm die bedarfsgewichtete Größe von Haushalten aus Österreich um durchschnittlich 1,1% pro Jahr zu, jene der Haushalte von außerhalb der EU15/EFTA Staaten aber um 3,0% (Tab. 8). Einem Zuwachsunterschied von rund 0,6 Prozentpunkten (Tab. 6; Rundungsdifferenz) beim Einkommen steht somit einer von rund 1,9 Prozentpunkten bei der Haushaltsgröße gegenüber, was in etwa das Dreifache ist. Das kann, unter anderem, ein Effekt einer unterschiedlichen Altersverteilung unter den Erwachsenen in den beiden Bevölkerungsteilen sein, was zu überprüfen wäre, bzw. können Familienzusammenführungen eine Rolle spielen.

Tab. 7: Bedarfsgewichtete Haushaltsgröße

Jahr	Herkunftsstaat			Zuzug nicht EU15/EFTA		
	AT	EU15/EFTA	Sonst	1960-1985	1986-1995	1996-2005
2008	1,7	1,5	1,9			
2009	1,7	1,6	1,9	1,5	1,8	1,8
2010	1,7	1,6	1,9	1,6	1,8	1,8
2011	1,7	1,6	1,9	1,5	1,8	1,9
2012	1,7	1,6	1,9	1,5	1,9	1,8
2013	1,7	1,7	1,9	1,5	1,8	1,7
Steigung	0,00	0,02	0,00	0,00	0,01	-0,01
%	-0,2	1,5	-0,2	-0,3	0,4	-0,4
n pro Jahr	Ca. 3500	130-165	525-560	57-84	159-178	80-89

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Tab. 8: Bedarfsgewichtete Haushaltsgröße

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	1,7	1,5	1,8	1,7	1,5	1,8
2	1,7	1,6	1,9	1,7	1,5	1,9
3	1,7	1,6	1,9	1,7	1,6	1,9
4	1,8	1,7	2,0	1,8	1,7	2,0
Steigung	0,02	0,05	0,06	0,02	0,07	0,05
%	1,1	3,2	3,0	1,3	4,2	2,6

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

4.3. Quellen des Wachstums des Haushaltseinkommens

Die prozentuell ähnliche Steigerung des Haushaltseinkommens kann sich aus unterschiedlichen Prozessen ergeben. Das jährliche Haushaltseinkommen nimmt zu, wenn entweder die Stundenlöhne der Haushaltsmitglieder steigen oder die Anzahl der bezahlten Stunden pro Woche oder der Beschäftigungswochen pro Jahr, wobei der Effekt auf das jährliche Haushaltseinkommen insofern begrenzt bleibt, weil dieses auch Einkommensersatzleistungen enthält und nicht nur die Erwerbseinkommen der Haushaltsmitglieder. Die Anzahl der bezahlten Stunden pro Woche nimmt zu, wenn entweder die beschäftigten Haushaltsmitglieder ihre Arbeitszeit verlängern oder zuvor nicht beschäftigte Haushaltsmitglieder neu in Beschäftigung kommen. Das Jahreseinkommen eines Haushalts entsteht somit aus der Multiplikation von Stundenlohn mit Wochenstunden pro beschäftigtem Haushaltsmitglied, mit Anzahl Beschäftigte im Haushalt und mit Beschäftigungswochen pro Jahr. Der Stundenlohn wird in SILC nicht erfragt, aber für die anderen drei Größen des Produkts und für das Ergebnis gibt es Daten. Die Beschäftigungswochen pro Jahr liegen nur un-

genau vor, aber für eine Einschätzung, wo die wesentlichen Quellen der Veränderungen der Jahreseinkommen der Haushalte liegen, reicht es.

Zugenommen hat vor allem die durchschnittliche Beschäftigtenzahl pro Haushalt. Sie stieg in den Haushalten mit Hauptverdienerin bzw. Hauptverdiener von außerhalb der EU15/EFTA pro Jahr um durchschnittlich 4,1% (Tab. 9). Die Steigerung der Beschäftigung entfiel zu fünf Achtel auf die weiblichen Haushaltsmitglieder. Die Zahl der Erwerbstätigen pro Haushalt war in Welle 4 gleich groß wie in den Haushalten mit Hauptverdienerin bzw. Hauptverdiener aus Österreich (1,4 Erwerbstätige). Bei den Haushalten aus Österreich war die Sachlage nicht gänzlich verschieden. Die durchschnittliche Beschäftigtenzahl pro Haushalt nahm aber nur um durchschnittlich 1,5% pro Jahr zu (nur zu drei Achteln weiblich) (Tab. 9).

Bei den durchschnittlichen wöchentlichen Arbeitsstunden pro Erwerbstätigen gab es Steigerungen von 0,4% pro Jahr in den Haushalten aus Österreich und 0,9% pro Jahr bei jenen von außerhalb der EU15/EFTA Staaten bzw. 1,3% und 1,1% in den Haushalten, die an allen vier Panelwellen teilnahmen (Tab. 10). Der Abstand sinkt von 2,4 Stunden in Welle 1 auf 1,5 Stunden in Welle 4.

Die jährlichen Beschäftigungsmonate im Haushalt nahmen ebenfalls zu und zwar mehr oder minder in Einklang mit der Zahl der Beschäftigten (Tab. 11), sodass sich daraus kein eigener Beitrag zur Steigerung des Haushaltseinkommens pro Jahr ergibt.

Tab. 9: Anzahl beschäftigter Haushaltsmitglieder

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	1,4	1,3	1,2	1,3	1,3	1,2
2	1,4	1,2	1,3	1,4	1,2	1,3
3	1,4	1,4	1,3	1,4	1,4	1,3
4	1,4	1,5	1,4	1,4	1,5	1,4
Steigung	0,02	0,08	0,05	0,03	0,08	0,05
%	1,5	5,9	4,1	2,4	6,2	3,7

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Tab. 10: Bezahlte wöchentliche Arbeitszeit pro beschäftigtem Haushaltsmitglied

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	33,7	36,4	31,3	32,8	36,5	31,3
2	33,7	35,4	32,2	33,3	33,1	32,2
3	33,9	37,9	30,9	33,7	37,8	31,8
4	34,1	35,8	32,6	34,1	35,8	32,6
Steigung	0,15	0,07	0,28	0,44	0,26	0,35
%	0,4	0,2	0,9	1,3	0,7	1,1

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Tab. 11: Beschäftigungsmonate im Haushalt im vorigen Kalenderjahr

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	16,8	15,5	15,2	16,4	15,7	14,9
2	16,9	16,4	15,5	16,6	16,3	15,8
3	17,1	16,4	15,4	17,1	16,6	16,0
4	17,3	17,6	17,3	17,3	17,6	17,3
Steigung	0,19	0,62	0,64	0,32	0,59	0,75
%	1,1	3,8	4,0	1,9	3,6	4,7

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

4.4. Fazit: Konvergenz?

Der Wachstumsvorsprung der eingewanderten Haushalte betrug bei der Beschäftigung somit 2,6 Prozentpunkte und bei der Arbeitszeit 0,5 Prozentpunkte. Das summiert sich auf einen Vorteil von 3,1 Prozentpunkten, um 2,5 Prozentpunkte mehr als die erwähnten 0,6 Prozentpunkte beim Haushaltsjahreseinkommen. Nun war die Rechnung hier nicht allzu genau, sodass man die 2,5 Prozentpunkte nur als groben Richtwert ansehen darf. Aber sie sind offenbar die Größenordnung, in der die Stundenlöhne der Mitglieder von Panelhaushalten aus Österreich rascher wuchsen als die Stundenlöhne von Mitgliedern von Panelhaushalten von außerhalb der EU15/EFTA Staaten. Während die Einkommenssteigerungen der Haushalte von außerhalb der EU15/EFTA Staaten somit vornehmlich auf mehr (weiblicher) Beschäftigung beruhten, entstanden sie in den Haushalten aus Österreich primär durch Zuwächse bei den Stundenlöhnen.

Das Einkommenswachstum erfolgte somit auf zwei gänzlich unterschiedliche Weisen. Man kann dasjenige der Haushalte von außerhalb den EU15/EFTA Staaten als extensiv bezeichnen, weil es durch Ausweitung der Erwerbstätigkeit erzielt wurde, jenes der Haushalte aus Österreich aber als intensiv, weil bei ihnen vorwiegend der Ertrag pro Arbeitsstunde gestiegen ist.

Wie diese Skizze zeigt, müsste man die Frage der Konvergenz der äquivalisierten Haushaltseinkommen differenzierter beantworten als die Fallzahlen in EU-SILC es ermöglichen. Im Großen ist der Befund einigermaßen klar: Weder in der Außensicht auf die Gesellschaft (Zeitverlauf) noch in der Innensicht der Haushalte (Panel) ereignete sich in den Jahren seit 2008 eine Angleichung der Einkommen zwischen eingewanderten und nicht eingewanderten Haushalten. Immerhin, die Einkommen der Haushalte mit mindestens einer Person im Erwerbsalter sind im Durchschnitt gestiegen, wenn auch nur wenig über der Inflation.

Wenn man das von der Frage der Konvergenz her betrachtet, dann ergibt sich aus den Panelverläufen das folgende Bild bezüglich der äquivalisierten Haushaltseinkommen und der fünf Faktoren, aus denen es entsteht:

- Divergenz bei den äquivalisierten Haushaltseinkommen
- Divergenz bei der bedarfsgewichteten Haushaltsgröße,
- Unklarer Trend bei den Haushaltseinkommen,
 - Unklarer Trend bei der wöchentlichen Arbeitszeit pro beschäftigtem Haushaltsmitglied,
 - Konvergenz bei der Beschäftigungsrate der Haushalte, besonders bei der weiblichen Beschäftigung,
 - Implizit Divergenz bei der Entlohnung pro Stunde.

Das heißt, im Zuge der Verlängerung der Aufenthaltsdauer gab es 2008 bis 2013 bei den zuvor schon in Österreich anwesenden Haushalten eine sichtliche Annäherung der Erwerbstätigkeit und besonders derjenigen der Frauen an das durchschnittliche Niveau der Haushalte aus Österreich und eine ebenfalls klar erkennbare Auseinanderentwicklung bei den bedarfsgewichteten Haushaltsgrößen sowie bei der durchschnittlichen Entlohnung pro Stunde, die aber gut zur beobachteten Steigerung der weiblichen Erwerbstätigkeit passt. Bezüglich der Entwicklung bei den nicht äquivalisierten Haushaltseinkommen wäre eine längere Beobachtungsreihe nötig, um zu eindeutigen Erkenntnissen zu gelangen. Gegenwärtig bewegen sie sich zu wenig zu jenen der Haushalte aus Österreich hin oder von ihnen weg, um darüber eine einigermaßen sichere Aussage treffen zu können.

5. Panelanalyse

5.1. Fragestellung

Die rein arithmetischen Überlegungen bis hierher zeigen, dass die Änderungen der äquivalisierten Haushaltseinkommen sich durch das Zusammenspiel einer Anzahl von Faktoren ergeben, das von Herkunftsgruppe zu Herkunftsgruppe verschieden ist, ohne dass einem Faktor allein dominante Bedeutung zukäme. Dennoch deuten die Indizien, so schwach sie sein mögen, auf eine wichtige Rolle des Einkommens pro Arbeitszeit hin, wenn es um die Erklärung von Unterschieden zwischen den Einkommen der Haushalte aus Österreich und jener von außerhalb der EU15/EFTA Staaten geht. Der Stundenlohn wird in EU-SILC nicht direkt abgefragt und ist auch aus Administrativdaten nicht eruierbar, aber man kann mit den Daten aus EU-SILC eine Ersatzgröße konstruieren. Sie bezieht das jährliche Nettoeinkommen eines Haushalts auf dessen Erwerbsintensität. Die Frage ist dann, ob die unterschiedlichen Bewegungen, denen diese Größe bei den einzelnen Herkunftsgruppen unterliegt, aufgeklärt werden können. Die Analyse von Paneldaten ist dabei, wie erörtert, grundsätzlich mit dem Anspruch verknüpft, kausale Erklärungen bereitstellen zu können. Die erste Frage ist folglich, welche Variablen einen Einfluss auf das Haushaltseinkommen nach dessen Bereinigung um die Erwerbsintensität des Haushalts haben können. In erster Linie ist in Österreich sicherlich an steigende Seniorität im Verbund mit kollektivvertraglichen Steigerungen zu denken. In Frage kommen aber auch Wechsel in besser bezahlte Tätigkeiten. Das können Branchenwechsel sein, aber auch Aufstiege. Mehr solcher Wechsel könnten eine raschere Steigerung der Einkommen nach sich ziehen. Von bloßen Arbeitgeberwechseln würde man das selbstverständlich nicht erwarten.

5.2. Die Erwerbsintensität des Haushalts

Die Beschäftigungsmonate pro Jahr und die Stundenzahl pro Woche der Haushaltsmitglieder lassen sich zu einem einzigen Indikator verdichten, der Erwerbsintensität des Haushalts. Die Variable liegt im SILC Datensatz in vier rangskalierten Ausprägungen vor, die bei Eurostat verbindlich so vereinbart wurden:

- Haushalte, bei denen die Variable nicht relevant ist
- Haushalte ohne Erwerbsbeteiligung
- Haushalte mit geringer Erwerbsbeteiligung (bis 20% des Potentials)
- Haushalte mit größerer oder voller Erwerbsbeteiligung (mehr als 20% des Potentials).

Sie wurde für die Zwecke dieser Untersuchung in eine intervallskalierte Variable mit den vier Werten 0, 0, 0,1 und 0,6 umkodiert. Leider ist im öffentlichen Datensatz keine eigene Ausprägung für hohe Erwerbsintensität enthalten, sondern sie ist mit der mittleren zusammengefasst, vermutlich aufgrund geringer Fallzahlen. Der Indikator ist in SILC mit Verwaltungsdaten für alle Jahre einheitlich kodiert.

Die Angabe zur Erwerbsintensität bezieht sich auf das Kalenderjahr, auf das sich auch die Angabe zum Haushaltseinkommen bezieht.

„Erwerbsintensität des Haushaltes: Weist den Anteil der Erwerbsmonate aller Personen zwischen 18 und 59 (ohne Studierende) an der maximal möglichen Erwerbszeit im Haushalt im Referenzjahr aus. Pro Person wird bei Vollzeit-erwerbstätigkeit unabhängig von den pro Monat tatsächlich geleisteten Stunden volle Erwerbsbeteiligung angenommen. Bei Teilzeiterwerbstätigkeit wird die aktuell geleistete Stundenzahl durch 35 dividiert und anteilmäßig eingerechnet. Ausgewiesen werden drei Kategorien: Keine oder sehr niedrige Erwerbsintensität: maximal 20% Erwerbsintensität, das entspricht in einem Einpersonenhaushalt einer ganzjährigen Erwerbstätigkeit von maximal sieben Wochenstunden. Mittlere Erwerbsintensität: mehr als 20% und weniger als 85% Erwerbsintensität im Haushalt, das entspricht in einem Einpersonenhaushalt ganzjährig einer Erwerbsbeteiligung von weniger als 30 Wochenstunden. Hohe Erwerbsintensität: 85% bis 100% Erwerbsintensität im Haushalt. Nicht ausgewiesen wird der Indikator für Personen über 59 Jahren und Personen in Haushalten ohne Personen im Erwerbsalter (hier: 18-59, ohne Studierende). Für den bis zur Berichterstattung 2008 ausgewiesenen Indikator „Erwerbsintensität des Haushaltes“ war das Erwerbsalter wie auch sonst im Bericht mit 20 bis 64 Jahren festgelegt. Die Berücksichtigung von Vollzeit- bzw. Teilzeiterwerbstätigkeit (Vollzeit=100%, Teilzeit=50%) war gegenüber der aktuellen Berechnung (Vollzeit=100%, Teilzeit entsprechend Stundenausmaß) weniger genau“ (Lamei u.a. 2015:23; siehe Atkinson/Marlier 2010).

Die Erwerbsintensität hat sich im Panelverlauf nur unmerklich verändert (Tab. 12). Das kann im ersten Moment bei den Haushalten von außerhalb der EU15/EFTA Staaten überraschen, weil bei ihnen die Beschäftigtenzahl pro Haushalt ja gestiegen ist. Offenbar ist bei ihnen auch die Zahl der

Personen in erwerbsfähigem Alter im selben Ausmaß gestiegen, sodass die Erwerbsintensität sich nicht bewegte.

Tab. 12: Erwerbsintensität im vorigen Kalenderjahr

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	0,5	0,5	0,5	0,5	0,5	0,5
2	0,4	0,5	0,5	0,4	0,5	0,5
3	0,4	0,5	0,5	0,4	0,5	0,5
4	0,4	0,5	0,5	0,4	0,5	0,5
Steigung	0,00	0,00	0,00	0,00	0,00	0,00
%	-0,7	0,0	0,1	-0,6	-0,3	-0,2

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

5.3. Auf die Erwerbsintensität abgeglichenes Nettohaushaltseinkommen

Das Nettohaushaltseinkommen dividiert durch die entsprechend gewichteten Kategorien der Erwerbsintensität des Haushalts ergibt eine mit der jährlichen Arbeitszeit gewichtete Version des Haushaltseinkommens pro Beschäftigten. Durch diesen Abgleich sollte das Einkommen an sich nicht mehr empfindlich sein für die sich ändernde Zahl der Erwerbstätigen im Haushalt und ihrer Wochenstunden.

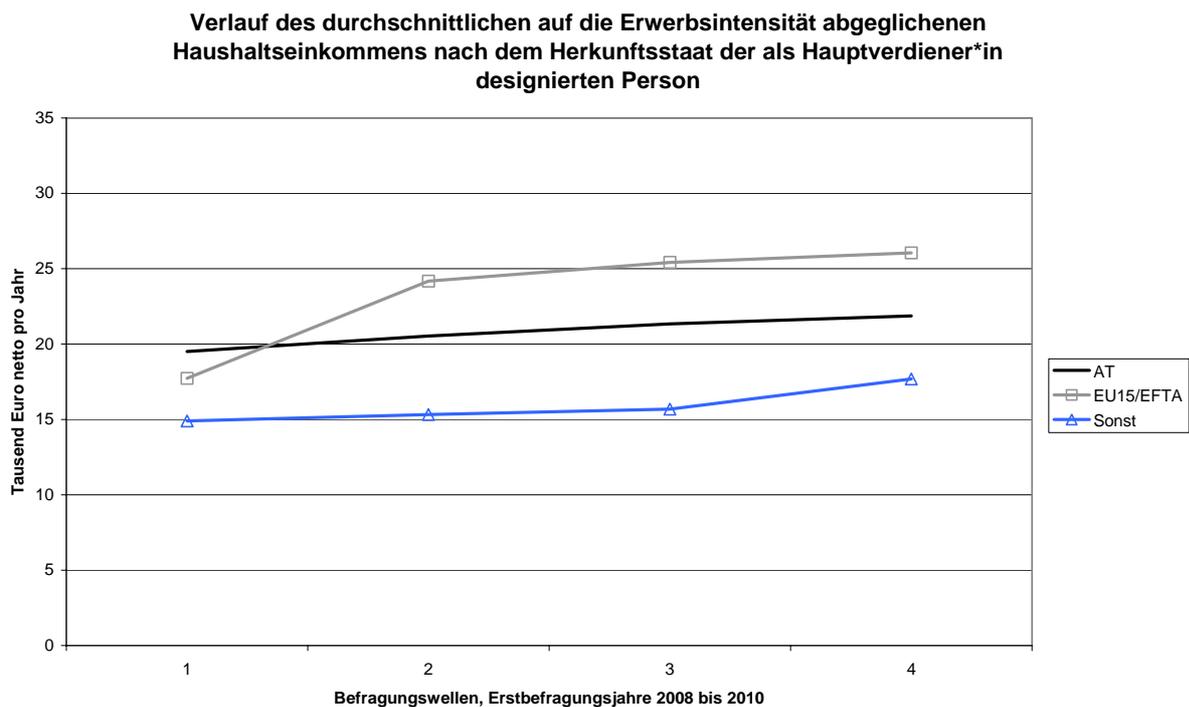
Die abgeglichenen Haushaltseinkommen bewegen sich im Vergleich zwischen Haushalten aus Österreich und solchen von außerhalb der EU15/EFTA Staaten auf sehr unterschiedlichen Niveaus, aber mit ähnlichen Steigerungen. Bei den Haushalten aus Österreich war von Welle 1 bis Welle 4 eine Steigerung von rund €19.500 auf rund €21.900 zu beobachten mit einer durchschnittlichen jährlichen Zunahme um €790, bei den Haushalten von außerhalb der EU15/EFTA Staaten eine Steigerung von rund €14.900 auf rund €17.700 und somit um jährlich €870 (Tab. 13). Deutlich rascher voran ging es bei den Haushalten aus den EU15/EFTA Staaten mit rasanten Steigerungen zu Beginn des Panels und Abflachung auf Zuwächse wie bei den anderen Herkunftsgruppen gegen Ende (Abb. 2). Der Befund ist ganz ähnlich, wenn nur jene Haushalte betrachtet werden, die an allen vier Panelwellen beteiligt waren. Es besteht somit Konvergenz zwischen den abgeglichenen Einkommen der Haushalte aus Österreich und aus Staaten außerhalb der EU15/EFTA, aber bei einer Lücke von etwa €4.200 und einer Annäherung von €80 pro Jahr dauert das Aufholen mehr als 50 Jahre.

Tab. 13: Nettohaushaltseinkommen pro Jahr abgeglichen auf die Erwerbsintensität, €000

Welle	Herkunftsstaat			Herkunftsstaat, alle 4 Wellen		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
1	19,5	17,7	14,9	19,6	18,5	15,6
2	20,5	24,2	15,3	20,5	23,2	15,5
3	21,3	25,4	15,7	21,6	24,7	16,6
4	21,9	26,0	17,7	21,9	26,0	17,7
Steigung	0,79	2,62	0,87	0,77	2,41	0,72
%	3,8	11,2	5,5	3,7	10,4	4,4

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Abb. 2



Die Veränderung des abgeglichenen Einkommens im Panelverlauf soll nun auf ursächliche Einflüsse hin analysiert werden.

5.4. Stichprobeneigenschaften

Bei Zeitreihendaten ist zu erwarten, dass aufeinander folgende Messungen bei der gleichen Person bzw. beim gleichen Haushalt eine gewisse Ähnlichkeit haben. Sie sind somit nicht voneinander unabhängig. Das Alter, etwa, ändert sich auf eine völlig vorhersehbare Weise, sodass sich jedes spätere Alter zweifelsfrei aus dem früheren herleiten lässt. Es gibt auch Merkmale, die sich gar nicht ändern, sondern für das ganze Leben festgelegt sind, und denen aus diesem Grund Sicherheitssysteme, seien sie staatlich oder privat, besonderes Interesse entgegenbringen. Doch auch veränderliche Merkmale, wie das Einkommen, haben eine Tendenz, wenn sie einmal hoch sind, hoch zu bleiben, und wenn sie einmal niedrig sind, niedrig zu bleiben. Diese Abhängigkeit späterer von früheren Messungen bedeutet, dass Regressionsverfahren, die grundsätzlich auf die Annahme unabhängig gemessener Werte bauen, um deren statistischen Fehler zu berechnen, eine Korrektur für Situationen brauchen, in denen die Unabhängigkeit der Messwerte nicht als gegeben angenommen werden kann. Die konkrete Korrektur hängt von der Struktur der Abhängigkeit ab. Es kann Situationen geben, in denen zwar unmittelbar aufeinander folgende Messungen eine Tendenz haben, ähnlich auszufallen, weiter zurückliegende aber nicht. Es gibt viele denkbare Muster für die Zusammenhänge je nach zeitlichem Abstand der Messungen zueinander. Diese Muster heißen Kovarianz- oder auch Korrelationsstrukturen und werden im Allgemeinen in Matrizenform dargestellt. Welchem Muster eine Stichprobe folgt, ist leider nicht ohne weiteres erkennbar und wird zudem durch das Modell beeinflusst, dessen Anpassung an die Daten man zu testen versucht. Man wird in der Wahl einer angenommenen Struktur theoretischen und empirischen Plausibilitäten folgen. Mitunter ist man aber auch durch die Fallzahl auf bestimmte Strukturen verwiesen, die auch dann noch die Konvergenz der Maximum Likelihood Schätzung zulassen, wenn die Zahl der Fälle (im Verhältnis zur Zahl der im Modell vorkommenden Variablen) nicht sehr groß ist. Das Ziel wird stets sein müssen, die einfachste, den Daten einigermaßen entsprechende Struktur zu finden (Twisk 2007:65), und daher möglicherweise so lange mit verschiedenen Korrelationsstrukturen zu experimentieren, bis man eine zufrieden stellende Lösung gefunden hat.

Ein Ausgangspunkt kann jedenfalls die Beobachtung der empirischen vorfindlichen Korrelationen in einer Stichprobe sein (Twisk 2007:65). Die Matrix der Korrelationen zwischen den Einkommen aus den vier Befragungswellen zeigt ein klares Muster, aber keines, das in gängiger Software als solches implementiert ist. Die Ergebnisse der Welle 1 korrelieren mit den drei Folgewellen, aber weniger als jene der Welle 2 mit den beiden folgenden und diese weniger als die Welle 3 mit der letzten (Tab. 14). Dies trifft sowohl in Bezug auf die äquivalisierten als auch auf die nicht äquivalisierten und die abgeglichenen Haushaltseinkommen zu. Die auf die Erwerbsintensität abgegliche-

nen Nettohaushaltseinkommen der Welle 2 korrelieren mit jenen der Welle 1 mit $r=0,77$, jene der Welle 3 mit der Welle 1 mit $r=0,72$ und jene der Welle 4 mit Welle 1 mit $r=0,68$. Je größer der Abstand zwischen den Wellen desto kleiner die Korrelation. Das gilt auch bei den Korrelationen mit der Welle 2: Die Welle 3 korreliert mit ihr mit $r=0,79$, die Welle 4 aber nur mit $r=0,75$. Beide Werte sind aber größer als dieselben mit entsprechenden Abständen bei der Welle 1. Zwischen Welle 3 und Welle 4, schließlich, beträgt die Korrelation $r=0,82$, also mehr als zwischen Welle 3 und Welle 2 ($0,78$) und zwischen Welle 2 und Welle 1 ($0,75$). Das abgegliche Einkommen korreliert auf den Abstand einer Welle etwas enger als das Einkommen per se, auf den Abstand von zwei oder drei Wellen aber immer weniger.

Tab. 14: Korrelationen des bedarfsgewichteten Nettohaushaltseinkommens pro Jahr (epinc, oberhalb der Diagonale) und des auf die Erwerbsintensität abgeglichenen Nettohaushaltseinkommens pro Jahr (hywi, unterhalb der Diagonale) zwischen den Befragungswellen, Stichprobenkohorten 2008 bis 2010, gewichtete Daten

hywi \ epinc	1	2	3	4
1	-	70	68	66
2	77	-	72	72
3	72	79	-	78
4	68	75	82	-

Quelle: Eigene Berechnungen anhand von EU-SILC Österreich.

Eine Verwandtschaft mit einer AR1-Struktur, bei der die Korrelationen umso kleiner sind je weiter das Feld von der Diagonale entfernt ist, ließe sich argumentieren. Man kann nicht ohne weiteres davon ausgehen, dass diese Matrix so bestehen bleibt, wenn das Einkommen zur abhängigen Variable in einer Regressionsanalyse wird, aber man hat einen Hinweis. Die beobachtete Struktur legt nahe, dass unter den implementierten Kovarianzstrukturen die unstrukturierte oder aber die AR1 zu wählen sei. Die unstrukturierte ist bezüglich der Fallzahlen anspruchsvoll, konvergiert in der Praxis aber dennoch oft leichter als die AR1.

Eine, wie hier, aus lediglich vier Messwerten bestehende Zeitreihe wird man nur unter Vorbehalten in die Zukunft oder auch in die Vergangenheit extrapolieren. Davon abgesehen wird in der Methodenliteratur aber auch laufend die Frage diskutiert, welche Fallzahlen eine Mehrebenenanalyse brauche, um gute Schätzwerte erbringen zu können. Ita Krefts nach wie vor diskutierte 30/30 Regel (Kreft/Yoon 1996) setzt sich bei einer Zeitreihenanalyse, wie hier, fast automatisch außer Kraft, weil man zwar auf der oberen Ebene schnell 30 Einheiten, seien es Personen, Haushalte oder Staaten, beisammen hat, auf der unteren Ebene aber kaum je 30 Messzeitpunkte aufweisen kann. Simulationsstudien seither weisen verschiedentlich in Richtung möglichst großer Zahlen auf der

oberen Ebene, aber nur geringer Zahlenerfordernisse auf der unteren, vor allem auch wenn auf der unteren Ebene alle Einheiten gleich viele Messwerte aufweisen (Pötschke 2014).

5.5. Wachstumskurvenanalyse

5.5.1. Ist eine komplexe Analyse notwendig?

Der Effekt der komplexeren Analyse zeigt sich in der Schätzung des Effekts des Zeitablaufs auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen und in der Präzision der Schätzung. Bei den Haushalten aus Österreich wird der Effekt des Zeitablaufs in einer linearen Regression mit €800 pro Jahr geschätzt. Im identischen, einfachen Modell wird er mit einer Mehrebenenanalyse mit Random Intercept auf €580 pro Jahr und mit Random Intercept und Random Slope auf €710 pro Jahr geschätzt (Tab. 15). Bei den Haushalten von außerhalb der EU15/EFTA Staaten reduziert sich der erwartete Effekt rein aufgrund dieser methodologischen Schritte von €820 auf €570 und €500 pro Jahr. Die Random Intercept Methode entfernt die Unterschiede zwischen den Haushalten aus der Schätzung des Zeiteffekts und belässt nur jenen der Zeit, wie sie sich innerhalb des jeweiligen Haushalts auswirkt. Im vorliegenden Fall führt das zu deutlichen Reduktionen des geschätzten Zeiteffekts.

Im Übergang von der linearen Regression zum Random Intercept Modell verringern sich auch die Standardfehler (Tab. 15). Dergleichen ist nur in einer Wachstumskurvenanalyse zu erwarten, also wenn die gemessenen Daten einen Zeitverlauf in sich bergen und dies in der Analyse ausdrücklich berücksichtigt wird. In Querschnittsanalysen erhöhen Mehrebenenmodelle den Standardfehler dagegen durchwegs (Twisk 2006:96).

Tab. 15: Der Effekt der Zeit auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen netto pro Jahr, €000 , gewichtet

	Koeffizient			Standardfehler		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Pooled OLS	0,80	2,83	0,82	0,01	0,04	0,01
Random Intercept	0,58	0,90	0,57	0,00	0,02	0,01
+ Random Slope	0,71	1,76	0,50	0,12	0,90	0,27

Durchgestrichene Werte sind nicht mit mindestens 95% Wahrscheinlichkeit von Null verschieden. Full Information Maximum Likelihood; Kovarianzstruktur Varianzkomponenten, bei Random Slope unstrukturiert. Separate Analyse für jedes der drei Herkunftsgebiete.

Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

Der konditionale Intraklassenkorrelationskoeffizient (ICC) zeigt in Form einer Korrelation an, wie stark die Werte auf der Zeitebene von jenen auf der Personenebene abhängig sind. Im vorliegenden Fall bewegen sich die drei ICC im Random Intercept Modell zwischen 0,82 und 0,85, sind also sehr hoch. Ein Mehrebenenmodell ist demnach hochgradig angezeigt, und eine einfache lineare Regression liefert, wie gezeigt (Tab. 15), stark verzerrte Koeffizienten für zeitgebundene Effekte (Hosoya u.a. 2014:193; Field 2009:729). In Längsschnittuntersuchungen ist stets mit einem hohen ICC zu rechnen, wenn auch nicht unbedingt mit einem so hohen (Twisk 2006:93).

5.5.2. Wechsel der beruflichen Tätigkeit

Ein bloßer Zeittrend ist keine Erklärung für die Veränderung der Einkommen, sondern beschreibt lediglich das, was zu erklären ist. Ziel muss es daher sein, den Koeffizienten auf der Zeitvariable wegzuerklären in dem Sinn, dass er Null wird, weil tatsächliche Gründe der Einkommensveränderungen in die Regressionsgleichung eingefügt worden sind. Um einem verbreiteten Missverständnis entgegenzutreten, es genügt nicht, ihn statistisch insignifikant werden zu lassen, denn das wäre auch durch eine kleine Fallzahl zu erreichen, sondern tatsächlich ihn in die Nähe von Null zu bekommen.

Um Einkommensverläufe zu erklären, ist es nahe liegend, auf Veränderungen der beruflichen Situation zu setzen. Im Rahmen von EU-SILC wird eine ganze Reihe von diesbezüglichen Informationen erhoben. Eine ist die ISCO-08 Berufsgruppe (Zweisteller), eine die berufliche Tätigkeit (37 Ausprägungen, die einerseits die arbeitsrechtliche Einordnung und andererseits die Qualifikationsanfordernisse der Tätigkeit spiegeln), die Befristung, die Ausübung einer Leitungsfunktion, relativ detailliert der Wirtschaftszweig des Betriebs, die Beschäftigtenzahl des Betriebs (genaue Angabe bis 10, dann in wenigen breiten Kategorien), die Anzahl der parallel ausgeübten Erwerbstätigkeiten, Wechsel des Arbeitsplatzes in den letzten 12 Monaten, der Grund des Arbeitsplatzwechsels (sieben Ausprägungen). Das sind durchwegs Angaben über Veränderungen bei den Beschäftigten im Haushalt, und die Häufigkeit solcher Veränderungen ist eine zeitveränderliche Variable, die sich für die Modellierung von Zeitverläufen eignet. Ausgewählt wurde unter den Genannten schließlich die berufliche Tätigkeit der Beschäftigten. Sie enthält Information über Auf- und Abstiege und über durchaus einkommensrelevante Wechsel zwischen den arbeitsrechtlichen Segmenten (Arbeiter, Angestellte, Vertragsbedienstete, Beamte, Selbständige). Sie weist zudem unter allen beschäftigungsbezogenen Variablen in SILC in den fraglichen Jahren die größte Häufigkeit von Änderungen auf. In 4.560 der 9.484 Haushalts-Jahr-Einheiten ab Welle 2 kommt mindestens ein Wechsel vor (48%).

Eine sehr einfache Wachstumskurvenanalyse des auf die Erwerbsintensität abgeglichenen Haushaltseinkommens (laut Verwaltungsdaten) der drei Stichprobenkohorten 2008, 2009 und 2010 über alle vier Wellen, in die außer der Zeitvariable (Welle) nur der Wechsel der beruflichen Tätigkeit durch beschäftigte Haushaltsmitglieder einbezogen ist, zeigt beim bloßen Zeitverlauf (Welle) eine Steigerung um €500 pro Jahr bei den Haushalten aus Österreich und eine fast dreimal so große bei jenen aus den EU15/EFTA Staaten, aber praktisch keine bei den Haushalten von außerhalb der EU15/EFTA Staaten. Bei allen drei ist das weniger als im Modell mit nur der Zeit und insbesondere bei den Haushalten von außerhalb der EU15/EFTA Staaten ist damit bereits eine sehr gute Annäherung an Null erreicht. Ein Teil der im einfachen Modell dem Wirken der Zeit zugeschriebenen Veränderung der Einkommen ergibt sich somit in Wahrheit aus den Tätigkeitswechseln.

Die Standardfehler zeigen, dass nur bei den österreichischen Haushalten die Wahrscheinlichkeit größer als 95% ist, dass die Steigerung nicht Null beträgt. Sie zeigen auch, dass die drei Steigerungen, so unterschiedlich sie anmuten mögen, statistisch nicht mit 95% Wahrscheinlichkeit verschieden sind. Die genannten Eurobeträge sind Mittelwerte über alle Haushalte der jeweiligen Herkunft. Unter den Haushalten einer jeden Herkunft gibt es aber bedeutsame Unterschiede in der rein durch den Zeitablauf bedingten Veränderung des Einkommens. Dies ist an den je nach Herkunft zwischen 46,15 und 67,49 betragenden Varianzen zu erkennen, wobei aber die Varianzen bei den Haushalten aus Österreich und bei jenen von außerhalb der EU15/EFTA Staaten praktisch gleich sind und auch jene bei den Haushalten aus EU15/EFTA Staaten nicht mit 95% Wahrscheinlichkeit von den beiden anderen verschieden ist. Das heißt, die Streuungen um die mittlere jährliche Einkommenssteigerung rein aufgrund des Zeitablaufs sind jedenfalls bei Österreich und bei außerhalb der EU15/EFTA Staaten gleich.

Pro Wechsel der beruflichen Tätigkeit von beschäftigten Haushaltsmitgliedern sind die Zuwächse bedeutend größer, nämlich rund €1.800 bei Haushalten aus Österreich, rund €2.900 bei Haushalten aus den EU15/EFTA Staaten und rund €2.200 bei Haushalten von außerhalb der EU15/EFTA Staaten. Da die Einkommen auf die Erwerbsintensität abgeglichen sind und in der Erwerbsintensität die Beschäftigungsmonate und die Arbeitszeit berücksichtigt sind, können die Effekte der Wechsel der beruflichen Tätigkeit nur in sehr beschränktem Maß durch gleichzeitige Erhöhung der Arbeitszeit oder durch Wechsel aus Saison- in Dauerbeschäftigung erklärt werden. Es muss sich, mit anderen Worten, weitgehend um tatsächliche Aufstiegseffekte handeln. Erneut sind jedoch die drei Eurobeträge nicht sicher verschieden, aber immerhin sind sie alle mit 95% Wahrscheinlichkeit nicht in Wirklichkeit Null. Die Varianz zwischen Haushalten des Effekts von Tätigkeitswechseln auf

den Einkommensverlauf ist mit Werten zwischen 129,3 und 199,8 erheblich. Sie ist zwischen den Haushalten von außerhalb der EU15/EFTA Staaten signifikant kleiner als zwischen den Haushalten aus Österreich. Die Haushalte aus EU15/EFTA Staaten liegen dazwischen, aber dass sich die Varianz bei ihnen signifikant von jener bei den beiden anderen Herkunftsgebieten unterscheidet, lässt sich nicht nachweisen.

Die Steigerungen muss man in Relation sehen zu den Ausgangsniveaus. Dem Modell nach – zu erkennen ist das an den Werten der Konstante – haben im Jahr vor der ersten Befragung die Haushalte aus Österreich netto €18.800 pro Erwerbseinheit verdient, jene aus den EU15/EFTA Staaten €19.400 und jene von außerhalb der EU15/EFTA Staaten €15.300. Den Wert für die EU15/EFTA Staaten darf man nicht allzu wörtlich nehmen, weil bei ihnen der Verlauf, wie man oben gesehen hat, durch einen linearen Zeitverlauf, wie er hier auch bei ihnen der Einfachheit halber unterstellt wurde, nicht adäquat abzubilden ist. Bei den anderen beiden Herkunftsniveaus ist der angenommene lineare Verlauf jedoch passend. Wenn nun der reine Zeitablauf den Haushalten aus Österreich €500 pro Jahr bringt, den Haushalten von außerhalb der EU15/EFTA Staaten aber nur €60, dann wird ihre Einkommensposition immer weiter auseinander driften. Die Lücke ließe sich durch vermehrte erfolgreiche Tätigkeitswechsel schließen, da jeder die Lücke um etwa €360 verringert. Das heißt aber, dass Haushalte von außerhalb der EU15/EFTA Staaten im Durchschnitt um mehr als einen Tätigkeitswechsel mehr pro Jahr zustande bringen müssten als Haushalte aus Österreich, um die jährliche Vergrößerung der Lücke rein aus dem Zeitablauf (€440) zu kompensieren. Wenn der Abstand zu den Haushalten aus Österreich auch noch verringert werden sollte, dann müssten es noch mehr Tätigkeitswechsel sein. Das ist unrealistisch.

Dass daneben der Zeitverlauf für Haushalte von außerhalb der EU15/EFTA Staaten praktisch nichts erbringt, obwohl ja auch sie an Kollektivvertragserhöhungen teilhaben, kann teils daher kommen, dass bei Haushalten, die von außerhalb der EU15/EFTA Staaten zugezogen sind, laufend Mitglieder neu in den Arbeitsmarkt eintreten, die infolge fehlender Vordienstzeiten und anderer Umstände wenig verdienen, sodass das auf die Erwerbsintensität des Haushalts abgeglichene Haushaltseinkommen nicht steigt. Ob dem so ist, müsste geprüft werden.

Die -2 Log Likelihood verringert sich gegenüber einem Modell, das neben der Konstanten nur die Zeit (Welle) enthält, substantiell und signifikant, nämlich um 36%, 23% und 31%. Das heißt, die Einkommensverläufe der Haushalte verdanken sich nicht einfach dem reinen Zeitablauf, sondern werden in bedeutendem Ausmaß durch aktive Chancennutzung der Haushalte am Arbeitsmarkt beeinflusst.

Tab. 16: Der Effekt der Zeit und des Wechsels der Tätigkeit von beschäftigten Haushaltsmitgliedern auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen netto pro Jahr, €000 , gewichtet

Mittel (Fixed)	Koeffizient			Standardfehler		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Konstante	18,80	19,40	15,03	0,33	1,78	0,70
Welle	0,50	1,46	0,06	0,13	0,76	0,33
Tätigkeit	1,83	2,92	2,19	0,30	1,34	0,62
Random: Varianzen um Mittel und Kovarianzen (Cov)						
Residuum	15,51	61,20	14,27	0,01	0,16	0,02
Konstante	389,47	467,71	289,63	9,31	57,45	18,07
Cov Konst-Welle	-61,64	-78,71	-76,28	2,78	19,22	7,31
Welle	47,43	67,49	46,15	1,25	9,63	3,36
Cov Konst-Tätigk	14,18	-23,74	43,42	5,88	40,83	12,08
Cov Welle-Tätigk	-50,44	-0,94	-52,31	2,22	16,31	5,26
Tätigkeit	199,81	143,24	129,33	6,05	26,31	10,88
Konditionaler ICC	0,96	0,88	0,95			
-2LL	37197467	2169022	7116447			
Verringerung	-21030580	-645417	-3230059			
N Haushalte un- gewichtet	3.707	142	640			

Durchgestrichene Werte sind nicht mit mindestens 95% Wahrscheinlichkeit von Null verschieden.
Separate Mehrebenenanalyse für jedes der drei Herkunftsgebiete.

Full Information Maximum Likelihood; Kovarianzstruktur unstrukturiert.

Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

Die Kovarianz zwischen der Konstante und dem Zeitverlauf ist negativ. Das besagt, dass bei den einzelnen Haushalten hohe Werte auf der Konstante mit niedrigen Koeffizienten auf der Zeit einhergehen und umgekehrt, und bedeutet, dass innerhalb jeder Herkunftsgruppe die Haushalte mit höherem abgeglichenem Anfangseinkommen im weiteren Verlauf schwächere Zuwächse durch den bloßen Zeitverlauf (Twisk 2006:93f, 97) haben als jene mit geringerem Anfangseinkommen. Anders ist es bei den Tätigkeitswechseln. Die Haushalte aus Österreich und von außerhalb der EU15/EFTA Staaten mit höheren Anfangseinkommen weisen durch Tätigkeitswechsel auch signifikant größere Zugewinne aus dem Zeitverlauf auf, während der Zusammenhang bei den Haushalten aus den EU15/EFTA Staaten insignifikant negativ ist. Die Kovarianz zwischen Zeit und Tätigkeitswechseln ist dagegen wieder bei allen drei Herkunftsnationen negativ, wenn auch bei EU15/EFTA in vernachlässigbarem Ausmaß. Das besagt, dass Haushalte mit größeren Zugewinnen aus dem bloßen Zeitverlauf kleinere aus den Tätigkeitswechseln haben und umgekehrt. Daraus lässt sich folgern, dass es einerseits Haushalte geben muss, die stärker über Tätigkeitswechsel Verände-

rungen des Einkommens erzielen, und andererseits Haushalte, die es auf anderem Weg tun, wenn wir auch erst herausfinden müssen, auf welchem.

5.5.3. Erwerbstätigkeit

Da die Haushaltseinkommen auf die Erwerbsintensität des Haushalts abgeglichen sind, sollte die Zahl der beschäftigten Haushaltsmitglieder an sich keine großen Auswirkungen auf das abgeglichene Einkommen haben. Dem ist aber nicht so. Wenn man die Zahl der Erwerbstätigen im Haushalt berücksichtigt, so erweist sie sich als erheblicher positiver Einfluss auf die abgeglichenen Einkommen. Ein zusätzliches erwerbstätiges Haushaltsmitglied erbringt demnach für Haushalte aus Österreich abgeglichen rund €7.000 netto pro Jahr, für Haushalte aus den EU15/EFTA Staaten rund €7.600 und für solche von außerhalb der EU15/EFTA Staaten rund €4.200. Die drei Effekte sind alle mit mehr als 95% Wahrscheinlichkeit größer als Null und ebenso der Unterschied zwischen Haushalten aus Österreich und solchen von außerhalb der EU15/EFTA Staaten.

Eine mögliche Erklärung für den Einkommenseffekt zusätzlicher Beschäftigter trotz des Abgleichs des Einkommens auf die Erwerbsintensität könnte sein, dass letztere nur in breiten Kategorien vorliegt, deren Gewichtung mit 0, 0,1 und 0,6 eventuell keine hinreichende Annäherung an die Realität ist. Möglicherweise müsste die obere Kategorie mit 0,7 oder 0,8 statt nur 0,6 gewichtet werden.

Die Koeffizienten auf der Zeitvariablen haben sich im Vergleich zum Modell ohne Änderungen der Erwerbshäufigkeit nicht signifikant verändert. Man beachte, dass nun auch der Effekt von Tätigkeitswechseln bei den Haushalten von außerhalb der EU15/EFTA Staaten kleiner ist als bei den Haushalten aus Österreich, wenn auch keineswegs in statistisch signifikanten Ausmaß. Zwar sind die Konstanten nun bei beiden Herkünften fast gleich groß, aber bei allen Effekten liegen die Haushalte aus Österreich im Vorteil.

Die -2 Log Likelihood verringert sich nochmals substantiell und signifikant, nämlich um 18%, 2% und 22%, sodass das ergänzte Modell deutlich aussagekräftiger ist als das vorherige.

Tab. 17: Der Effekt der Zeit, des Ausmaßes der Erwerbsbeteiligung und des Wechsels der Tätigkeit von beschäftigten Haushaltsmitgliedern auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen netto pro Jahr, €000, gewichtet

Mittel (Fixed)	Koeffizient			Standardfehler		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Konstante	10,65	8,56	11,08	0,69	2,68	1,27
Welle	0,31	1,74	-0,12	0,15	0,67	0,36
Erwerbstätige	6,95	7,59	4,25	0,42	1,76	0,68
Tätigkeit	1,85	3,03	1,75	0,35	1,35	0,68
Random: Varianzen um Mittel und Kovarianzen (Cov)						
Residuum	5,73	53,42	4,34	0,00	0,14	0,01
Konstante	1187,12	766,84	755,04	33,08	142,62	52,18
Cov Konst-Welle	-98,74	-156,47	-131,51	5,70	28,76	12,15
Welle	64,62	53,70	52,94	1,73	7,44	3,75
Cov Konst-Erwerb	-481,45	-358,78	-260,59	16,72	81,80	23,05
Cov Welle-Erwerb	0,40	71,87	33,63	3,03	17,43	5,67
Erwerbstätige	331,10	269,21	141,16	10,59	51,34	12,08
Cov Konst-Tätig	-45,52	-24,72	164,98	11,69	58,56	19,91
Cov Welle-Tätig	-67,12	4,54	-61,01	2,96	14,08	5,84
Cov Erwerb-Tätig	26,10	6,46	-59,40	6,80	40,36	10,64
Tätigkeit	276,81	140,36	146,28	8,25	25,05	12,32
Konditionaler ICC	1,00	0,93	0,99			
Cov Konst-Erwerb	-481,45	-358,78	-260,59	16,72	81,80	23,05
Cov Welle-Erwerb	0,40	71,87	33,63	3,03	17,43	5,67
Erwerbstätige	331,10	269,21	141,16	10,59	51,34	12,08
Cov Konst-Tätig	-15,52	-24,72	164,98	11,69	58,56	19,91
Cov Welle-Tätig	-67,12	4,51	-61,01	2,96	14,08	5,84
Cov Erwerb-Tätig	26,10	6,46	-59,40	6,80	40,36	10,64
Tätigkeit	276,81	140,36	146,28	8,25	25,05	12,32
Konditionaler ICC	1,00	0,93	0,99			
-2LL	30588948	2126831	5581087			
Verringerung	-6608538	-42192	-1535381			
N Haushalte ungewichtet	3.707	142	640			

Durchgestrichene Werte sind nicht mit mindestens 95% Wahrscheinlichkeit von Null verschieden.
Separate Mehrebenenanalyse für jedes der drei Herkunftsgebiete.
Full Information Maximum Likelihood; Kovarianzstruktur unstrukturiert.
Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

Die Kovarianzen zwischen der Konstante und dem Zeitverlauf sowie der Erwerbstätigkeit sind alle negativ. Dagegen ist die Kovarianz zwischen Erwerbstätigkeit und Wechsel der Tätigkeit ist bei den Haushalten aus Österreich signifikant positiv, bei den Haushalten aus den EU15/EFTA Staaten insignifikant positiv und bei den Haushalten von außerhalb der EU15/EFTA Staaten signifikant negativ. Bei den ersteren haben somit Haushalte, die aus zusätzlicher Beschäftigung trotz Ab-

gleichs auf die Erwerbsintensität Einkommenszugewinne erzielen, der Tendenz nach gleichzeitig auch Einkommenszugewinne aus Tätigkeitswechseln, während bei Herkunft von außerhalb der EU15/EFTA Staaten Haushalte mit größeren Zugewinnen aus zusätzlicher Beschäftigung kleinere aus Tätigkeitswechseln haben und umgekehrt.

Jene Haushalte aus Österreich und von außerhalb der EU15/EFTA Staaten, die größere Zugewinne durch den bloßen Zeitablauf erleben, haben geringere durch Tätigkeitswechsel und umgekehrt. Bei den Haushalten aus den EU15/EFTA Staaten gibt es keinen diesbezüglichen Zusammenhang.

Unter den Haushalten, die nicht aus Österreich sind, haben jene mit größeren Zugewinnen rein durch den Zeitablauf auch größere durch zusätzliche Beschäftigung. Bei den Haushalten aus Österreich gibt es keinen derartigen Zusammenhang.

5.5.4. Bildung

Die Bildung von erwachsenen Individuen ist in Längsschnittdaten häufig als feststehend behandelt worden. Man könnte versucht sein, sie in der Terminologie von Giesselmann & Windzio (2014:102) als quasi-zeitkonstant zu bezeichnen, was aber an sich nicht richtig wäre, weil die Variable ja nicht konzeptionell zeitkonstant ist und empirisch gestört wird, sondern umgekehrt, sie ist konzeptionell zeitabhängig, erweist sich aber empirisch als wenig veränderlich. Auf die durchschnittliche Bildung in Haushalten trifft weder das eine noch das andere zu. Sie ist vor allem durch die Alterung der Kinder sowie durch Ein- und Auszüge von Haushaltsmitgliedern konzeptionell und empirisch veränderlich. Verwendet wurden im vorliegenden Fall nur die höchsten Bildungsabschlüsse von Personen im Erwerbsalter, die nicht in Ausbildung stehen. Diese Begrenzungen wurden getroffen, um auf die beschäftigungsrelevante Bildung zu fokussieren. Sie bringen allerdings auch zusätzliche Variabilität in die Daten, über deren Wünschbarkeit man diskutieren könnte. In 3.715 der 9.484 Haushalts-Jahr-Einheiten ab Welle 2 kommt mindestens ein Wechsel vor (39%). Die Bildung wurde auf ihren Gesamtdurchschnitt zentriert.

Die Ergebnisse verändern sich gegenüber dem vorangehenden Modell im Detail, aber nicht in der Struktur. Der Zeiteffekt bleibt bei den Haushalten aus Österreich mit rund €370±320 statistisch signifikant von Null verschieden. Die Hinzufügung der Bildung hat hier keine Annäherung an Null oder an Insignifikanz gebracht, sodass weitere Überlegungen anzustellen wären, welche Ursachen in Frage kämen. Bei den Haushalten von außerhalb der EU15/EFTA Staaten ist der Zeiteffekt nun fast genau Null. Das gilt aber auch für den Effekt der Bildung. Während eine Steigerung der durchschnittlichen Höhe des Abschlusses um eine Schulstufe den Haushalten aus Österreich eine Ein-

kommenssteigerung pro Jahr von etwa €1.280±900 bringt, sind es bei Haushalten von außerhalb der EU15/EFTA Staaten €30±100. Der fatale Eindruck aus anderen Studien, dass Bildung eingewanderten Haushalten nichts bringt, verstärkt sich neuerlich. Tätigkeitswechsel und Steigerung der Erwerbstätigkeit bleiben die einzigen Mittel für sie, um das Einkommen zu erhöhen.

Bei den Haushalten aus EU15/EFTA Staaten gelang es der Maximum Likelihood Prozedur nicht mehr, zu konvergieren. Hier ist die Fallzahl selbst mit Haushaltsgewichten bereits zu klein.

Tab. 18: Der Effekt der Zeit, des Ausmaßes der Erwerbsbeteiligung und des Wechsels der Tätigkeit von beschäftigten Haushaltsmitgliedern auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen netto pro Jahr, €000, gewichtet

Mittel (Fixed)	Koeffizient			Standardfehler		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Konstante	9,27		11,54	1,22		1,44
Welle	0,37		-0,02	0,16		0,36
Erwerbstätige	5,22		3,43	0,52		0,87
Tätigkeit	1,46		1,49	0,36		0,69
Bildung	1,28		-0,03	0,45		0,51
Random: Varianzen um Mittel und Kovarianzen (Cov)						
Residuum	3,02		2,80	0,00		0,00
Konstante	3183,86		643,36	115,52		69,52
Cov Konst-Welle	-98,92		-62,51	12,08		12,36
Welle	69,36		45,95	1,94		3,63
Cov Konst-Erwerb	-385,21		-68,28	31,55		25,13
Cov Welle-Erwerb	-5,61		17,55	4,19		6,57
Erwerbstätige	401,98		188,56	16,36		20,86
Cov Konst-Tätig	60,60		105,25	23,06		24,84
Cov Welle-Tätig	-77,01		-54,07	3,42		5,97
Cov Erwerb-Tätig	6,34		-51,36	9,38		14,63
Tätigkeit	266,61		139,39	8,91		13,15
Cov Konst-Bild.	-973,23		-150,20	33,86		17,74
Cov Welle-Bild.	23,05		-12,38	4,21		3,73
Cov Erwerb-Bild.	-13,99		-8,60	10,79		9,14
Cov Tätigk.-Bild.	-31,33		-8,18	8,33		7,99
Bildung	453,76		92,36	13,12		7,05
Konditionaler ICC	1,00		1,00			
-2LL	26347541	1321369	5017880			
Verringerung	-4241408	-805462	-563206			
N Haushalte ungewichtet	3.707	142	640			

Durchgestrichene Werte sind nicht mit mindestens 95% Wahrscheinlichkeit von Null verschieden.

Separate Mehrebenenanalyse für jedes der drei Herkunftsgebiete.

Full Information Maximum Likelihood; Kovarianzstruktur unstrukturiert.

Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

Twisks Warnung vor dem falschen Glauben, die Wahl der Kovarianzstruktur sei nebensächlich (Twisk 2007:64f), bestätigt sich teilweise. Mit AR1 statt unstrukturiert wird der Zeiteffekt bei den Haushalten aus Österreich kleiner und der Standardfehler größer, sodass er statistisch insignifikant wird (Tab. 19). Die Vergrößerung des Standardfehlers, die auch bei den Haushalten von außerhalb der EU15/EFTA Staaten auftritt, ist aber zugleich ein Hinweis, dass unstrukturiert die bessere Wahl ist als AR1. Auch der Bildungskoeffizient bei den Haushalten aus Österreich ist mit AR1 viel kleiner als mit unstrukturiert, während der Standardfehler praktisch gleich bleibt, im Verhältnis zum Koeffizienten also viel größer wird. Bei den Haushalten von außerhalb der EU15/EFTA Staaten ist beides größer. Auch fast alle anderen Standardfehler sind mit AR1 größer als mit unstrukturiert. Bei den Haushalten aus den EU15/EFTA Staaten konvergiert die Maximum Likelihood Prozedur auch weiterhin nicht.

Die -2 Log Likelihood ist nur marginal verschieden von der Version mit unstrukturiert.

Tab. 19: Der Effekt der Zeit, des Ausmaßes der Erwerbsbeteiligung und des Wechsels der Tätigkeit von beschäftigten Haushaltsmitgliedern auf das auf die Erwerbsintensität abgeglichene Haushaltseinkommen netto pro Jahr, €000, gewichtet

Mittel (Fixed)	Koeffizient			Standardfehler		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Konstante	14,55		12,54	0,77		1,03
Welle	0,21		-0,10	0,41		0,61
Erwerbstätige	3,06		1,86	0,57		0,84
Tätigkeit	1,27		0,38	0,51		0,82
Bildung	-0,15		0,19	0,44		0,62
Random						
Residuum	3,02		2,80	0,00		0,00
AR1 diagonal	479,12		153,01	6,79		5,51
AR1 rho	-0,05		-0,07	0,01		0,03
Konditionaler ICC	0,99	1,00	0,98			
-2LL	26357948	1322734	5018896			
N Haushalte ungewichtet	3.707	142	640			

Durchgestrichene Werte sind nicht mit mindestens 95% Wahrscheinlichkeit von Null verschieden.

Separate Mehrebenenanalyse für jedes der drei Herkunftsgebiete.

Full Information Maximum Likelihood; Kovarianzstruktur AR1.

Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

5.6. Fazit

Die Ausgangsfrage war eine im Wesentlichen methodologische, nämlich ob es gelinge, die Zeitverläufe im SILC Panel mit seinen vier Messzeitpunkten zu modellieren und zu erklären. In Hinblick darauf sind nun auch Schlussfolgerungen zu ziehen. Das inhaltliche Vehikel dazu waren die Haushaltseinkommen im Zeitraum 2009 bis 2013 bzw. eigentlich 2008 bis 2012. In Bezug auf sie können auch inhaltliche Schlüsse gezogen werden. Wenn wir diese vorwegnehmen, dann zeigt die Untersuchung, dass Einkommenssteigerungen im Zeitverlauf bei Haushalten aus Österreich nicht aus den gleichen Gegebenheiten entspringen wie bei Haushalten von außerhalb der EU15/EFTA Staaten. Nur Tätigkeitswechsel haben bei beiden in etwa den gleichen Effekt. Zusätzliche Beschäftigung hat in den Haushalten aus Österreich den Anzeichen nach im Durchschnitt größere Einkommenswirksamkeit, wenngleich die statistische Signifikanz dieser Aussage unter 90% liegt. Bildungszunahmen erweisen sich in Haushalten von außerhalb der EU15/EFTA Staaten als wirkungslos, jedenfalls kurzfristig, während sie in Haushalten aus Österreich sofort das Einkommen erhöhen. Schließlich gibt es in den Haushalten aus Österreich auch noch einkommenssteigernde Faktoren, deren Natur hier nicht identifiziert werden konnte, die aber in den Haushalten von außerhalb der EU15/EFTA Staaten fehlen.

Konvergenz der abgeglichenen Einkommen ereignete sich in den letzten Jahren offenbar sehr langsam oder gar nicht. Wenn sie sich ereignet, dann vermutlich primär im Generationswechsel, also nicht innerhalb von Haushalten, sondern in Form ihrer Neugründung. Eine Datengrundlage, um diesen Effekt mit mehr als Indizien zu belegen, dürfte in Österreich aber bisher nicht existieren.

Man muss sich im Klaren sein, dass diese Generalisierungen den Durchschnitt breiter Gruppen von Haushalten beschreiben. Es gibt in jeder der drei Gruppen Haushalte mit größeren und mit kleineren Einkommenszuwächsen. Dieser Punkt wäre bei Gelegenheit noch zu vertiefen.

Was nun die methodologische Frage betrifft, so ging es darum, die Eignung der Daten für panelspezifische statistische Verfahren zu prüfen. Die sechs oben nacheinander präsentierten Modelle unterscheiden sich durch zunehmende Komplexität des Modells und der Analysemethode. Dass der statistische Aufwand seine Berechtigung und seinen Nutzen hat, zeigt sich am Verlauf der Werte der $-2 \text{ Log Likelihood}$ ($-2LL$). Sie verringern sich sukzessive (Tab. 19). Am Schluss liegen sie bei allen drei Herkunftsgebieten unter der Hälfte des Anfangswerts. Ein sehr großer Teil der Verringerung ereignet sich dabei im Wechsel von linearer Regression (Pooled OLS) zu Mehrebenenanalyse mit Random Intercept beim einfachsten Modell. Bei den Haushalten aus den EU15/EFTA Staaten ist der Anteil an der Gesamtverringerng durch diesen einen, rein methodolo-

gischen Schritt sogar noch bedeutend größer. Geht man noch einen methodologischen Schritt weiter und lässt beim einfachsten Modell nicht nur Random Intercept, sondern zusätzlich auch Random Slope zu, so ist bei zwei der drei Herkünfte bereits mehr als die Hälfte der Verringerung der -2LL erreicht. Die Hinzufügung von drei erklärenden Variablen setzt den Prozess fort. (Die EU15/EFTA Angabe im letzten Modell wird berichtet, obwohl das Modell nicht konvergierte).

Fast selbstredend sind diese Verringerungen der -2 Log Likelihood alle weit größer als die kritischen Werte der Chi-Quadrat-Verteilung bei der begleitenden Zunahme der Freiheitsgrade für 95% oder auch 99% Wahrscheinlichkeit der Verschiedenheit von Null verlangen. Jedes der Modelle erbringt damit einen statistisch höchst signifikanten Zugewinn an Erklärungskraft gegenüber dem vorangehenden.

Tab. 20: Modellanpassung anhand der -2 Log Likelihood (-2LL), gewichtet

	-2LL			Prozent		
	AT	EU15/EFTA	Sonst	AT	EU15/EFTA	Sonst
Pooled OLS	58228047	2814439	10346506			
Random Intercept	45817481	2320604	8342517	78,7	82,5	80,6
+ Random Slope	40912293	2225936	7610362	70,3	79,1	73,6
mit Tätigkeit	37197467	2169022	7116447	63,9	77,1	68,8
mit Erwerb	30588948	2126831	5581087	52,5	75,6	53,9
mit Bildung	26347541	1321369	5017880	45,2	46,9	48,5
Änderung						
RI-OLS	-12410566	-493835	-2003990	-21,3	-17,5	-19,4
RIRS-RI	-4905188	-94668	-732155	-10,7	-4,1	-8,8
mit Tätigkeit	-3714826	-56914	-493915	-9,1	-2,6	-6,5
mit Erwerb	-6608518	-42192	-1535361	-17,8	-1,9	-21,6
mit Bildung	-4241408	-805462	-563206	-13,9	-37,9	-10,1
Freiheitsgrade				Änderung	kritischer Wert	
Pooled OLS	3	3	3	Freiheitsg.	5%	1%
Random Intercept	4	4	4	1	3,841	6,635
+ Random Slope	6	6	6	2	5,991	9,210
mit Tätigkeit	10	10	10	4	9,488	13,277
mit Erwerb	15	15	15	5	11,070	15,086
mit Bildung	21	21	21	6	12,592	16,812

Full Information Maximum Likelihood; Kovarianzstruktur unstrukturiert.

Separate Mehrebenenanalyse für jedes der drei Herkunftsgebiete.

Daten: Stichprobenkohorten 2008 bis 2010, EU-SILC 2008 bis 2013 mit Verwaltungsdaten, bereitgestellt von Statistik Austria.

Dennoch darf man sich über die Erklärungsmächtigkeit der Analyse keine Illusionen machen. Eine einfache lineare Regression, die die hierarchische Datenstruktur ignoriert, ergibt ein korrigiertes R-

Quadrat von 0,37 bei den Haushalten aus Österreich und 0,26 bei jenen von außerhalb der EU15/EFTA Staaten. Das verweist auf eingehenderen Forschungsbedarf als in dieser Exploration der Möglichkeiten, die sich mit EU-SILC bieten, zu leisten war, und möglicherweise auf die Notwendigkeit eines Panels mit mehr als drei Jahren Laufzeit sowie auf die potentielle Bedeutsamkeit von Variablen, die sich nicht sprunghaft, sondern kleinweise und kontinuierlich ändern.

Ein längeres Panel wäre besonders auch dann nötig, wenn sich Änderungen nicht sofort auswirken, sondern ihre Wirkung mit Verzögerung oder auch allmählich über einen längeren Zeitraum entfalten. Das könnte, etwa, bei betrieblichen und außerbetrieblichen Weiterbildungsmaßnahmen so sein. Der Engpass ist jedoch, dass in EU-SILC darüber nur sehr unsichere Information vorliegt, sodass ein längeres Panel nicht helfen würde.

Kontinuierliche Änderungen können gerade bei Einwanderinnen und Einwanderern bedeutsam sein. Sie sind vielfach darauf verwiesen, ihre Deutschkenntnisse und zumal ihre Dialektkenntnisse unstrukturiert im Alltag zu erwerben. Dabei ist zwar anzunehmen, dass im Allgemeinen nach relativ kurzer Zeit ein Plafond erreicht wird, einfach weil das (positive) Feedback fehlt. Es wird aber dennoch eine größere Zahl von Personen geben, bei denen der Deutscherwerb durch Selbststudium, besonderes Interesse oder besondere sprachliche Begabung kontinuierlich weitergeht. Ein Panel kann zeigen, wo, wann und wie sich das auswirkt. Die Kontinuität des Vorgangs macht es aber unmöglich, nach konkreten Ereignissen zu fragen, in denen sich der Deutscherwerb plastisch ausdrückt. Man müsste hier andere und teils vielleicht auch in der Umfrageforschung wenig erprobte Wege gehen. Angesichts des hohen Stellenwerts in der öffentlichen Diskussion, wären Fakten selbstverständlich von enormer Wichtigkeit, sodass die Investition an sich gerechtfertigt wäre. Freilich, von der ersten Absicht über die Präzisierung des Bedarfs und die Entwicklung der geeigneten Erhebungsmethoden bis zu Daten aus einem mehrjährigen Panel kann ein Jahrzehnt auch vergehen. Ein Gegenargument ist das allerdings nicht, denn es ist nicht zu erwarten, dass die Frage des Deutscherwerbs und seiner Wirkungen in zehn Jahren obsolet geworden sein wird.

Der eingehendere Forschungsbedarf ist zweifelsfrei gegeben. Es wurden hier nur wenige Variablen auf ihre Wirksamkeit hin untersucht. Weitere erklärende Variablen hinzuzufügen hat seine Schwierigkeiten. Es kommt nicht von ungefähr, dass alle gerechneten Lehrbuchbeispiele von Mehrebenenanalysen nur ein Minimum an Variablen enthalten, denn Mehrebenenmodelle sind in der Praxis schwer dazu zu bringen, zu konvergieren (Field 2009:776), und die Ursachen für die ausbleibende Konvergenz zu finden bzw. das Problem auf eine statistisch vertretbare Art und Weise zu lösen, ist bisher mehr eine Kunst als eine Wissenschaft, denn die Lehrbücher schweigen dazu, wie selten zu etwas. Zum Beispiel, könnte man die Zeit auch quadratisch in das Modell auf-

nehmen, wie es im vorliegenden Fall bei den Haushalten aus EU15/EFTA Staaten geboten wäre. Das bereitet als solches keine Schwierigkeiten. Lehrbuchbeispiele zeigen gerne den Effekt eines quadratischen und eines kubischen Terms der Zeitvariable. Es gibt aber kein einziges gerechnetes Lehrbuchbeispiel, in dem der quadratische Term mit Random Slope vorkommt, und der Grund ist ganz einfach, dass ein Mehrebenenmodell unter dieser Bedingung nicht konvergiert, aus welchen Gründen auch immer (Twisk 2006:97f, 104).

Eine Kunst auszuüben aber erfordert Übung, sehr viel Übung. Dieser Befund verweist auf zwei Erfordernisse. Das eine ist, dass ein sachkundiger Umgang mit den nicht ganz, aber fast grenzenlosen Möglichkeiten von Mehrebenenmodellen nicht auf Knopfdruck erwartet werden kann, sondern über einen längeren Zeitraum kontinuierlich neue Herausforderungen verlangt. Das zweite Erfordernis ist, dass in erheblichem Maß Zeit für den Anschluss an die akademische Welt sein muss, in der die Transformation von Erfahrungskönnen in leistungsstarke nachvollziehbare Prozeduren erfolgt. Mit anderen Worten, der Bedarf an angewandter Forschung ist im Bereich der Analyse von Paneldaten ohne Spielraum für Grundlagenforschung gegenwärtig nicht zu decken.

6. Quellenangaben

Atkinson, Anthony B / Marlier, Eric (2010) Statistics on Income and Living Conditions; Eurostat.

Brüderl, Josef (2010) Kausalanalyse mit Paneldaten; in: Wolf/Best (Hg) 2010:963-994.

Field, Andy (2009) Discovering Statistics Using SPSS, third edition; Sage.

Giesselmann, Marco / Windzio, Michael (2012) Regressionsmodelle zur Analyse von Paneldaten; VS.

Giesselmann, Marco / Windzio, Michael (2014) Paneldaten in der Soziologie: Fixed Effects Paradigma und empirische Praxis in Panelregression und Ereignisanalyse; Kölner Zeitschrift für Soziologie und Sozialpsychologie 66/1:95-113.

Hagenaars, Aldi J.M. / Vos, Klaas de / Zaidi, M. Ashgar (1994) Poverty Statistics in the Late 1980s: Research Based on Micro-data; Office for Official Publications of the European Communities
http://bookshop.europa.eu/en/poverty-statistics-in-the-late-1980s-pbCA8394636/?pgid=y8dIS7GUWMdSR0EAIMEUUsWb000032P4Mn1;sid=MmmmZuucB4ymZbt-WAk4wYm5fwq4FKb_LBs=?CatalogCategoryID=82cKABstbqMAAAEjs5EY4e5L, 2013-01-17.

Hosoya, Georg / Koch, Tobias / Eid, Michael (2014) Längsschnittdaten und Mehrebenenanalyse; Kölner Zeitschrift für Soziologie und Sozialpsychologie 66/Supp:189-218.

Kreft, Ita G.G. / Yoon, Bokhee (1996) Are Multilevel Techniques Necessary? An Attempt at Demystification <http://files.eric.ed.gov/fulltext/ED371033.pdf>, 2015-08-26.

Lamei, Nadja / Glaser, Thomas / Heuberger, Richard / Oismüller, Anneliese / Riegler, Romana (2015) Tabellenband EU-SILC 2014 Einkommen, Armut und Lebensbedingungen; Statistik Austria.

Langer, Wolfgang (2009) Mehrebenenanalyse. Eine Einführung für Forschung und Praxis, 2. Auflage; Verlag für Sozialwissenschaften.

Langer, Wolfgang (2010) Mehrebenenanalyse mit Querschnittsdaten; in: Best/Wolf (Hg) 2010:741-774.

Pötschke, Manuela (2014) Aktuelle Probleme der Modellierung von Mehrebenen-Daten; Kölner Zeitschrift für Soziologie und Sozialpsychologie 66/Supp:219-239.

Twisk, Jos W. R. (2006) Applied Multilevel Analysis; Cambridge University Press.

Twisk, Jos W. R. (2007) Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide; fourth printing with corrections; Cambridge University Press.

Wolf, Christof / Best, Henning (Hg) (2010) Handbuch der sozialwissenschaftlichen Datenanalyse; VS Verlag für Sozialwissenschaften.

Wooldridge, Jeffrey M. (2010) Econometric Analysis of Cross Section and Panel Data, 2nd Edition; MIT Press.

Wooldridge, Jeffrey M. (2013) Introductory Econometrics: A Modern Approach, 5th Edition; Cengage.