



Arbeitskreis Quantitative Steuerlehre
Quantitative Research in Taxation – Discussion Papers

Malte Chirvi

**Arbeiten Frauen aufgrund des Ehegattensplittings
weniger?**

Ein quasi-experimenteller Ansatz für Deutschland

arqus Discussion Paper No. 217

May 2017

www.arqus.info

ISSN 1861-8944

Arbeiten Frauen aufgrund des Ehegattensplittings weniger?

Ein quasi-experimenteller Ansatz für Deutschland

Malte Chirvi

Humboldt-Universität zu Berlin

Zusammenfassung

Die vorliegende Arbeit soll einen Beitrag zur Klärung der Frage leisten, ob das Ehegattensplitting, wie in vielen Arbeiten behauptet, einen negativen Einfluss auf das Arbeitsverhalten verheirateter Frauen hat. Hierzu wird für Deutschland erstmals ein quasi-experimenteller Ansatz gewählt, bei dem anhand des allgemeinen Effekts der Heirat auf das Erwerbsverhalten der heiratenden Frauen die Wirkung des Ehegattensplittings identifiziert wird. Die Veränderung des Arbeitsverhaltens heiratender Frauen wird dabei der Veränderung bei jenen Frauen gegenübergestellt, die im gleichen Zeitraum weiterhin mit ihrem festen Partner zusammenleben, ohne zu heiraten (Kontrollgruppe). Die Ergebnisse sprechen eindeutig gegen einen unmittelbaren negativen Effekt des Ehegattensplittings auf das Erwerbsverhalten von Frauen. Jedoch reduzieren heiratende und zeitgleich ein Kind bekommende Frauen ihre Erwerbstätigkeit stärker als gebärende Frauen in der Kontrollgruppe.

JEL-Classification: D10, H24, H31, J12, J22

Gliederung

1	Einleitung.....	3
2	Literaturüberblick.....	4
3	Das Ansatz dieser Arbeit	10
4	Daten.....	13
5	Deskriptive Auswertung.....	15
6	Methodik	21
7	Ergebnisse	25
7.1	Coarsened Exact Matching.....	25
7.2	Validierung der Ergebnisse	31
8	Schlussfolgerungen.....	32
9	Literatur	35
10	Anhang.....	39

1 Einleitung

Es gibt wenige Bestandteile des deutschen Einkommensteuerrechts, die in den letzten Jahrzehnten so sehr Teil der öffentlichen Diskussion waren, wie die gemeinsame Besteuerung von Ehepaaren (§ 26b EStG).¹ Dabei wird das zu versteuernde Einkommen beider Ehepartner zusammengerechnet und anschließend bei jedem Partner die Hälfte dem Tarif des § 32a Abs. 1 EStG unterworfen (Ehegattensplitting).² Die gemeinsame Steuer entspricht also im Ergebnis der Steuer, die jeder der beiden Ehepartner bei einem gemittelten zu versteuernden Einkommen zu tragen hätte. Aufgrund des progressiven Verlaufs des Einkommensteuertarifes ergibt sich so bei unterschiedlichen Einkommen der Ehepartner in den meisten Konstellationen eine im Gegensatz zur Einzelveranlagung geringere Steuerbelastung. Diese Regelung wird häufig als eine Benachteiligung von verheirateten Frauen als der im derzeitigen Erwerbsgefüge typischen Gruppe von Zweitverdienern³ gesehen, wobei ihre Wirkung an folgendem Beispiel erläutert sei:

Das zu versteuernde Einkommen des Ehemannes liegt über 107.322,- €⁴, während die Ehefrau keinerlei Einkünfte erzielt. Sollte die Ehefrau nun überlegen, einer Erwerbstätigkeit nachzugehen, läge der Eingangssteuersatz für „ihre“⁵ Einkünfte bei der Zusammenveranlagung bereits bei 42%⁶, während sie bei Individualbesteuerung den Grundfreibetrag und die Progressionszone ausnutzen könnte und so bis zu einem Einkommen von 53.666,- € unterhalb dieses Grenzsteuersatzes läge. Das Ehegattensplitting, so die Behauptung, beeinflusse die Ehefrau folglich hinsichtlich ihrer Entscheidung zu arbeiten. In diesem Zusammenhang lohnt auch ein Blick auf die Entwicklung des Netto- bzw. verfügbaren Einkommens abhängig von der Arbeitsentscheidung des Zweitverdieners⁷, der sich der Arbeit von Wrohlich (2007) entnehmen lässt.⁸ So betrachtet sie beispielhaft Ehepaare mit einem bzw. zwei Kindern und unterschiedlich hohen Löhnen. In den

¹ Dabei steht nicht immer nur die Abschaffung des Ehegattensplittings, z. B. aufgrund vermeintlicher Arbeitsmarkteffekte, zur Debatte. National wurde das Ehegattensplitting bspw. durch ein Urteil des BVerfG v. 07.05.2013 (2 BvR 1981/06) auf eingetragene Lebenspartnerschaften ausgeweitet. International gibt es z. B. in Kanada eine gegensätzliche Entwicklung und ein (beschränktes) Ehegattensplitting wurde nach langer Diskussion zum Jahr 2015 eingeführt; siehe <http://www.cra-arc.gc.ca/gncy/bdgt/2014/qa10-eng.html>. Zur vorausgehenden Diskussion siehe bspw. Krzepkowski/Mintz (2013).

² Siehe hierzu § 32a Abs. 5 EStG.

³ Diese Aufteilung innerhalb der Ehe, die bis zum 1. EheRG aus dem Jahr 1976 noch gesetzlich kodifiziert war – die Ehefrau durfte nur dann einer Erwerbstätigkeit nachgehen, wenn es mit ihren Pflichten in Ehe und Familie vereinbar war – ist nach Baumgarten/Houben (2014) tendenziell auch aktuell zu beobachten.

⁴ Ab einem gemeinsamen zu versteuernden Einkommen in dieser Höhe gilt für Ehepaare im Veranlagungszeitraum 2016 ein Grenzsteuersatz von 42%.

⁵ Dies gilt natürlich für jegliche zusätzlichen Einkünfte eines Ehepaares, unabhängig davon, wer sie erwirtschaftet.

⁶ Sofern keine geringfügige Beschäftigung aufgenommen wird.

⁷ Der Zweitverdiener ist hier immer die Ehefrau, für die bei diesen Berechnungen auch niedrigere (reale) Löhne zu Grunde gelegt werden. Bei gleichen Löhnen wäre der dargestellte Nettoeinkommenszuwachs bzw. der Zuwachs des verfügbaren Einkommens dementsprechend etwas größer.

⁸ Siehe S. 58 – 63.

verschiedenen Konstellationen steigt das verfügbare Einkommen des Haushalts durch eine zusätzliche Vollzeitarbeit der Ehefrau zwischen 7% (Löhne in Höhe des 25sten Perzentils der Gesamtverteilung; ein einjähriges Kind) und 41% (Medianlöhne; ein vierjähriges Kind). In jedem Fall ist der Einkommenszuwachs, wenig überraschend, weit von einer Verdopplung entfernt. Neben dem Ehegattensplitting spielen hier aber vor allem auch die entstehenden Kosten der Kinderbetreuung sowie die Beiträge zur Sozialversicherung eine große Rolle. Von daher ist nicht verwunderlich, dass die OECD (2015) Deutschland eine der höchsten „participation tax rates“ für Zweitverdiener innerhalb der OECD-Staaten attestiert. Nur Dänemark und Island weisen hier in sämtlichen betrachteten Einkommensgruppen höhere Steuersätze auf.⁹

In Zusammenhang mit dem Ehegattensplitting wird regelmäßig auch auf die im internationalen Vergleich geringe Erwerbsbeteiligung verheirateter Frauen in Deutschland hingewiesen¹⁰, wengleich aktuelle Daten des Statistischen Amtes der Europäischen Union diese Behauptung nicht stützen. So lag die Erwerbsbeteiligung verheirateter Frauen in Deutschland im Jahr 2013 mit 70,1% deutlich über dem EU28-Durchschnitt mit 62,2% und EU-weit auf Platz sieben hinter vor allem skandinavischen und baltischen Staaten. Die Erwerbsbeteiligung verheirateter Männer in Deutschland war allerdings deutlich höher (86,8%).¹¹

Das Ziel dieser Arbeit ist es, zur Beantwortung der Frage beizutragen, ob das Ehegattensplitting verheiratete Frauen in Deutschland tatsächlich davon abhält, zu arbeiten. Der hierzu gewählte Ansatz und die Notwendigkeit eines weiteren Beitrags zu diesem Themenbereich werden in den nächsten Kapiteln erläutert.

2 Literaturüberblick

Das Ehegattensplitting wurde und wird nicht nur in den Wirtschaftswissenschaften, hier insbesondere in der Finanzwissenschaft und der Betriebswirtschaftlichen Steuerlehre, sondern auch in angrenzenden Disziplinen wie der Steuerrechtswissenschaft umfassend diskutiert.¹² Seit

⁹ Siehe Abschnitt PF 1.4: Neutrality of tax-benefit systems. Die „participation tax rate“ beinhaltet dabei sowohl die Einkommensteuer als auch Sozialversicherungsbeiträge. Betrachtet wird der Fall eines verheirateten Paares mit zwei Kindern (4 und 6 Jahre), in dem das Einkommen des Erstverdieners bei 67% des Durchschnittseinkommens liegt und der in den Arbeitsmarkt eintretende Partner alternativ 50%, 67% bzw. 100% des Durchschnittseinkommens verdient.

¹⁰ Siehe z. B. Bach et al (2011), S. 13; Fehr/Kallweit/Kindermann (2015), S. 53. Wagenhals (2008), S. 254.

¹¹ Diese Ausführungen sind der Fußnote 152 von Maiterth/Chirvi (2015) entnommen und beruhen auf eigener Auswertung der Eurostat-Daten. Gleichzeitig zeigt der Report von Cornelißen (2005), beruhend auf Daten aus dem Jahr 2004, dass die Erwerbstätigenquote von Frauen mit zwei und mehr Kindern deutlich unterhalb des OECD-Durchschnitts und den meisten europäischen Ländern sowie der Quote von Männern liegt (Grafik 5.4).

¹² Eine Darstellung sämtlicher Facetten dieser Diskussion soll nicht Teil dieser Arbeit sein, Überblicke mit vielen weiterführenden Literaturhinweisen finden sich beispielsweise bei Sausgruber/ Winner (1996) und Maiterth/Chirvi (2015).

geraumer Zeit wird die Diskussion jedoch von den Arbeitsangebotseffekten des Ehegattensplittings dominiert. Hierbei richtet sich das Augenmerk auf die Arbeitsangebotselastizität. Die meisten Analysen – so z. B. Bargain/Orsini/Peichl (2014) mit einem internationalen Überblick – finden höhere Arbeitsangebotselastizitäten bei verheirateten Frauen im Vergleich zu verheirateten Männern. Ausgehend von der Optimalsteuertheorie wird daher gefolgert, dass unter Effizienzgesichtspunkten eine Besteuerung mit niedrigeren Steuersätzen für Ehefrauen (im Vergleich zu ihren Ehepartnern) angezeigt sei.¹³ Die Höhe der geschätzten Elastizitäten schwankt allerdings aufgrund unterschiedlicher Datengrundlagen und angewandter Methodik deutlich, so dass eine genaue Einschätzung der Arbeitsangebotseffekte nicht möglich ist. Außerdem scheint die Vorstellung homogener Elastizitäten innerhalb der beiden Gruppen unrealistisch. In der folgenden Übersicht sollen Arbeiten betrachtet werden, die konkretere empirische Erkenntnisse zur Frage, inwiefern das Arbeitsverhalten verheirateter Frauen durch das Ehegattensplitting beeinflusst wird, leisten können.

Betrachtet man empirische Arbeiten zur Wirkung des Ehegattensplittings in Deutschland, lassen sich noch Anfang der 1990er Jahre¹⁴ ausschließlich Analysen basierend auf Modellen, in denen das Arbeitsangebot des Ehemannes als fix angenommen wird, finden.¹⁵ In den letzten zehn bis fünfzehn Jahren haben sich dann für Deutschland Analysen basierend auf diskreten Haushaltsnutzenmodellen mit flexibler Berücksichtigung beider Partner durchgesetzt. Die meisten mikroökonomischen Analysen beziehen sich dabei auf das unitäre strukturelle Arbeitsangebotsmodell¹⁶ aus der Arbeit von van Soest (1995). Es wird davon ausgegangen, dass Ehepartner (ohne die Berücksichtigung anderer Haushaltsmitglieder) ihre Arbeitsstunden gemeinsam nutzenmaximierend abstimmen. Die tatsächlich beobachteten Arbeitsstunden werden daher, abhängig von persönlichen Charakteristika und Präferenzen für Konsum und Freizeit bzw. Arbeit, als die nutzenmaximierenden angesehen. Die auf dieser Basis geschätzten Präferenzen dienen im nächsten Schritt dazu, diverse Steuerreformvorschläge zum Ehegattensplitting zu evaluieren: Welche Arbeitsangebotskombination würde ein Paar mit den geschätzten Präferenzparametern wählen, wenn sich das aus dem Arbeitseinkommen

¹³ Siehe z. B. bei Boskin/Sheshinski (1983), Apps/Rees (2011) oder Alesina/Ichino/Karabarbounis (2011), die sogar eine zusätzliche Differenzierung nach Familienstand in Betracht ziehen. Schröder/Burow (2016) finden dagegen in einem Experiment das Gegenteil heraus: Männer als Zweitverdiener reagieren auf Steuersatzänderungen, Frauen dagegen nicht.

¹⁴ Siehe bspw. Gustafsson (1992) und Spahn/Kaiser/Kassella (1992).

¹⁵ Diese Modelle werden auch „male chauvinist models“ genannt.

¹⁶ Neben den unitären Modellen („unitary models“) gibt es auch einzelne Arbeiten, die ihre Analysen auf Basis kollektiver Modelle („collective models“) durchführen, in denen nicht ein einheitlicher Haushaltsnutzen maximiert wird, sondern das Arbeits- und Konsumverhalten sämtlicher Haushaltsmitglieder berücksichtigt wird. Für weitere Informationen siehe bspw. bei Beninger/Laisney/Beblo (2007).

resultierende Haushaltsbudget bzw. die Grenzbelastung dieser Einkommen aufgrund einer Steuerreform, z.B. bei Abschaffung des Ehegattensplittings, ändern würde?

So untersuchen auf Basis dieses modellgestützten Ansatzes neben anderen Steiner/Wrohlich (2004), Wrohlich (2007), Bach et al. (2011), Müller et al. (2013) und Bonin et al. (2013) Reformen des Ehegattensplittings in Deutschland und kommen allesamt zu dem Ergebnis, dass sich sowohl das extensive als auch das intensive Arbeitsangebot verheirateter Frauen bei Abschaffung des Ehegattensplittings zugunsten einer Individualbesteuerung erhöhen würde. Die Ergebnisse sind in ihrer Höhe unterschiedlich, liegen bei der prozentualen Änderung der Partizipationsrate (extensives Arbeitsangebot) jeweils im unteren einstelligen, bei der prozentualen Änderung der Arbeitsstunden (intensives Arbeitsangebot) dagegen teilweise sogar im zweistelligen Bereich. Umgerechnet in Vollzeitäquivalente finden die Arbeiten eine Erhöhung des Arbeitsangebots verheirateter Frauen von bis zu 243.000 Vollzeitäquivalenten¹⁷ bei gleichzeitiger geringfügiger Senkung¹⁸ des Arbeitsangebots der Ehepartner. Zum Vergleich: Insgesamt betrug das Arbeitsvolumen in Deutschland im Zeitraum der aufgezählten Untersuchungen immer über 30 Mio. Vollzeitäquivalente.¹⁹ Bei einer Modifizierung des Ehegattensplittings hin zum sogenannten Realsplitting²⁰ sind die geschätzten Effekte allerdings deutlich geringer.²¹

Die Schätzung von Arbeitsangebotseffekten auf Basis struktureller Modellierung, wie sie in den vorgenannten Arbeiten vollzogen wird, ist trotz ihrer Verbreitung in der Literatur nicht frei von Schwächen. So spielt, wie in entsprechenden Arbeiten auch erwähnt, das Einkommen anderer Haushaltsmitglieder keine Rolle in den Modellen.²² Außerdem werden bei der Modellierung der (Haushalts)Nutzenfunktionen zwangsläufig Annahmen getroffen²³, welche Faktoren bei der Nutzenmaximierung und Budgetberechnung der Haushalte eine Rolle spielen. Neben der Berücksichtigung von Freizeit und Konsum als (einzige) nutzengenerierende Faktoren betrifft dies auch Annahmen bzgl. weiterer Details wie bspw. der Kinderbetreuung, „fixer Kosten der Erwerbsbeteiligung“ oder Haushaltsarbeit.

¹⁷ Siehe Bonin et al. (2013).

¹⁸ Bei Bonin et al. (2013) reduziert sich das Arbeitsangebot der Ehemänner um 20.000 Vollzeitäquivalente.

¹⁹ Siehe z.B. bei Asef/Wanger/Zapf (2011), S. 1063.

²⁰ Dies bezeichnet den auf maximal 13.805,- € (seit dem Veranlagungszeitraum 2002) beschränkten Übertrag eines nachehelichen Unterhalts bzw. eines Trennungunterhalts von dem leistenden (durch Sonderausgabenabzug) auf den empfangenden (als Hinzurechnung auf sein Einkommen) Ex-Partner.

²¹ Auf Untersuchungen einer Ausweitung des Ehegattensplittings zu einem sogenannten Familiensplitting, z. B. nach dem Vorbild Frankreichs, bei dem Einkommen steuerlich auch auf Kinder „verteilt“ werden kann, soll in dieser Arbeit nicht weiter eingegangen werden. Arbeiten hierzu sind z. B. die von Wrohlich (2007), und Steiner/Wrohlich (2008).

²² Siehe Steiner/Wrohlich 2004, S. 8.

²³ Zutreffend ist allerdings der Hinweis von Keane (2010): „One is forced to accept that all empirical work in economics, whether “experimentalist” or “structural”, relies critically on a priori theoretical assumptions“. Die einzige Ausnahme mögen hier vollkommen randomisierte Experimente sein. Auch in den weiteren Ausführungen zu dieser Arbeit wird deutlich werden, dass zur Identifikation der Ergebnisse einige weitreichende Annahmen getroffen werden.

Außerdem wird in den Arbeiten implizit davon ausgegangen, dass die beobachteten Individuen die genauen steuerlichen Folgen ihrer Arbeitsentscheidungen überblicken und aus diesem Wissen rationale Schlussfolgerungen ziehen. Arbeiten wie z. B. von Blaufus et al. (2015) zeigen allerdings, dass eine solche steuerliche Einschätzung den Steuerpflichtigen oftmals nicht gelingt. Entscheidungen mögen daher eher einer „gefühlten Belastung“, z. B. aufgrund der Wahl der Steuerklassenkombination III/V bei der Lohnsteuer, folgen.

Die umfangreiche theoretische Fundierung struktureller Ansätze mag zwar der Identifikation kausaler Effekte förderlich sein – und ist gleichzeitig alternativlos hinsichtlich ex-ante zu evaluierender Reformszenarien – allerdings werden in den genannten Arbeiten lediglich Korrelationen im Status Quo abgebildet, aus denen nicht ohne Weiteres Rückschlüsse auf in den untersuchten Reformszenarien bestehende kausale Zusammenhänge abgeleitet werden können. Wenn die zu Grunde liegenden theoretischen Modelle die Zusammenhänge fehlerhaft abbilden, dürften die geschätzten Präferenzen und die auf dieser Basis ermittelten Auswirkungen von Reformen falsch sein.

Zur Identifizierung kausaler Effekte haben sich insbesondere in der Arbeitsmarktökonomie aber auch generell in den Gesellschaftswissenschaften in den letzten zwei Jahrzehnten (quasi-) experimentbasierte Analysen etabliert, während strukturelle Ansätze an Bedeutung verlieren.²⁴ Die allgemeine Frage, inwiefern eine vorherige strukturelle Modellierung in empirischen Untersuchungen grundsätzlich zweitrangig ist, wie es bspw. Angrist/Krueger (1999) in ihrem Überblick zu empirischer Arbeitsmarktökonomie im Sinne einer „reduced-form“-Lehre suggerieren, oder unentbehrlich ist, wie bspw. von Keane (2010) behauptet, soll hier nicht geklärt werden.²⁵ Ein gut designtes Quasi-Experiment kann nach Meinung des Autors als ex-post-Analyse²⁶ eine wertvolle Alternative zu strukturellen Modellen²⁷ sein und eigene Erkenntnisse liefern, die – so sieht es auch Keane – zur Validierung der Ergebnisse struktureller Modelle beitragen können.

²⁴ Vgl. z.B. Rust (2010), S. 21, der diese Aussage sowohl für die Forschung als auch die Lehre an führenden amerikanischen Universitäten trifft. Er verweist in seinem Beitrag auch auf die seiner Meinung nach negative Entwicklung der gesamten Wirtschaftswissenschaft, in der der theoretischen Fundierung eine immer geringere Bedeutung zukommt. Beispielhaft erwähnt er den im Jahr 2003 mit der John Bates Clark Medal ausgezeichneten Ökonomen Steve Levitt, der nach eigener Aussage weder gut in Mathe und Ökonometrie sei, noch sich in ökonomischer Theorie auskenne und trotzdem einen Superstar-Status unter den Ökonomen hätte.

²⁵ Eine umfangreichere Arbeit zu dieser vielschichtigen Diskussion kommt von Wolpin (2013), der eher als Verfechter zweitgenannter Denkrichtung gelten mag.

²⁶ Strukturelle Modelle eignen sich offensichtlich besser für ex-ante Analysen, obwohl in dieser Arbeit ebenfalls versucht wird, auf Basis von ex-post Ereignissen eine potentielle Steuerreform ex-ante einzuschätzen.

²⁷ Trotzdem sollten ex-post durchgeführte Quasi-Experimente natürlich immer auf theoretischen Überlegungen basieren.

Mangels geeigneter zu untersuchender Reformen des Ehegattensplittings gibt es in Deutschland bislang keine, aber auch international nur wenige Studien, die einen quasi-experimentellen Ansatz zur Untersuchung von Arbeitsverhaltensreaktionen durch eine gemeinsame Besteuerung von Ehepartnern wählen. Hier sind insbesondere die Arbeiten von Eissa (2002), Crossley/Jeon (2007), LaLumia (2008) und Selin (2014) zu erwähnen.

Eissa (2002) schätzt in ihrer Arbeit, inwiefern verheiratete Frauen ihr Arbeitsangebot bei sinkenden Steuersätzen durch eine Steuerreform (den „Tax Reform Act of 1986“) in den USA anpassen, wobei sie in ihrer Untersuchung einen Differenz-in-Differenzen-Ansatz nutzt. Als Treatment- und Kontrollgruppen wählt sie Frauen, die jeweils unterschiedlich stark von reformbedingten Änderungen ihrer Grenzsteuerbelastungen betroffen sind. Zur Einteilung in die beiden Gruppen dient dabei – aufgrund möglicher Endogenitätsprobleme bei Identifizierung über das eigene Einkommen – die Höhe des nicht von der Ehefrau erwirtschafteten Haushaltseinkommens. Die Ergebnisse dieses quasi-experimentellen Ansatzes bestätigen die Ergebnisse früherer Schätzungen, nämlich hohe Arbeitsangebotselastizitäten verheirateter Frauen in den USA in den 1980er Jahren. Allerdings basieren die Ergebnisse auf der doch sehr restriktive Annahme, wonach das Arbeitsverhalten der Ehefrau nicht vom Arbeitsverhalten anderer Haushaltsmitglieder beeinflusst wird, vice versa.

Crossley/Jeon (2007) untersuchen, inwiefern die Reduktion gemeinsamer Besteuerung in Folge des Ersatzes eines in Kanada bis 1988 gültigen, von der Höhe der Einkünfte des Zweitverdieners abhängigen Steuerabzugs beim Erstverdiener (höhere Einkünfte des Zweitverdieners führten zu einer Besteuerung in gleicher Höhe zum Grenzsteuersatz des Erstverdieners) hin zu einer Steuergutschrift beim Erstverdiener, die sich bei höheren Einkünften des Zweitverdieners lediglich in Höhe des Eingangssteuersatz reduzierte. Die Autoren nutzen einen Differenz-in-Differenzen-Ansatz, wobei die betrachteten Frauen mit maximal mittlerer Bildung²⁸ anhand der Einkommen der Ehemänner in Treatment- und Kontrollgruppe eingeteilt werden²⁹, da die Grenzbelastung des Arbeitslohns verheirateter Frauen mit einkommensstarken Ehepartnern (Treatmentgruppe) durch die betrachtete Reform deutlich stärker sinkt als bei Frauen mit einkommensschwächeren Ehepartnern (Kontrollgruppe). Die Autoren kommen zu dem Ergebnis, dass diejenigen Frauen, die mit besser verdienenden Männern verheiratet waren, ihr extensives Arbeitsangebot nach der Steuerreform im Vergleich zur Kontrollgruppe signifikant erhöhten und daher tatsächlich auf die Ausdehnung einer gemeinsamen Besteuerung reagieren.

²⁸ Es werden nur Frauen betrachtet, die maximal einen High-School-Abschluss haben.

²⁹ Die Einteilung – Frauen mit Männern im 85. bis 99. Einkommensperzentil zählen zur Treatment-, Frauen mit Männern im 21. bis 35. Einkommensperzentil zur Kontrollgruppe – scheint recht optimistisch im Sinne einer möglichst zufallsähnlichen Gruppeneinteilung.

Eine zur Beantwortung der im vorliegenden Beitrag interessierenden Frage – reduzieren verheiratete Frauen ihr Arbeitsangebot aufgrund des Ehegattensplittings oder nicht? – ideales natürliches Experiment nutzt LaLumia (2008) in ihrer Studie. Vor 1948 gab es in neun US-Bundesstaaten eine gemeinsame Ehegattenbesteuerung (sog. „community property states“), wohingegen in den restlichen US-Bundesstaaten (sog. „common law states“) eine Individualbesteuerung vorherrschte. Durch den Revenue Act 1948 wurde eine gemeinsame Besteuerung von Ehegatten auf Bundesebene eingeführt, die aber nur in den „common law states“ eine Steueränderung bewirkte. Durch die Verwendung der „community property states“ als Kontrollgruppe und (Befragungs-)Daten aus den Jahren 1940 und 1950 schätzt die Autorin verschiedene Auswirkungen der eingeführten gemeinsamen Besteuerung auf Arbeitsmarktvariablen. Trotz Anstiegen sowohl beim extensiven (Arbeitsmarktpartizipation in der Woche vor der Befragung; Wert 1940: 18,5%, Wert 1950: 29,5%) als auch beim intensiven Arbeitsangebot (Arbeitsstunden pro Woche; Wert 1940: 9 Stunden, Wert 1950: 13,5 Stunden) verheirateter Frauen in den common law states, identifiziert die Autorin eine Reduzierung (gemessen durch einen im Vergleich weniger starken Anstieg) beider Werte infolge der Steuerreform. Aufgrund der Einführung der gemeinsamen Besteuerung in den common law states sank ersterer um etwa 2,2 Prozentpunkte und letzterer um etwa 1,1 Stunden, vor allem getrieben durch die gesunkene Partizipationsrate. Beide Schätzungen sind signifikant, beziehen sich allerdings vor allem auf Entwicklungen von Frauen mit relativ hoher Bildung. Außerdem ergibt die Analyse, dass der Anteil verheirateter Frauen, die Nicht-Arbeitseinkommen erwirtschaften, nach Einführung der gemeinsamen Besteuerung abgenommen hat. Dies fasst die Autorin nachvollziehbar als Indiz für eine zurückgegangene Einkommensübertragung zwischen den Ehepartnern im Falle gemeinsamer Besteuerung auf.

Ebenfalls auf eine tatsächliche Reform der Ehegattenbesteuerung bezieht sich die Arbeit von Selin (2014). Der Autor untersucht Veränderungen der Arbeitsmarktpartizipation verheirateter Frauen vor und nach der Abschaffung der gemeinsamen Besteuerung von Ehepartnern im Jahr 1971 in Schweden. Er differenziert diese nach dem Einkommen des Partners bzw. dem eigenen Grenzsteuersatz der Ehefrau bei der Entscheidung, zu arbeiten. Er zeigt, dass die Zunahme der Arbeitsmarktpartizipation nach der Reform umso größer war, desto stärker der Grenzsteuersatz der Ehefrau durch diese gesenkt wurde (= je höher der Grenzsteuersatz vor der Reform war). Größer sind die Effekte außerdem bei Frauen mit Kindern, was der Autor auch auf die besonders niedrige Arbeitsmarktpartizipation dieser Frauen vor der Reform zurückführt. Da die geschlechterspezifische Zusammensetzung der Arbeitskräfte in Schweden im beobachteten Zeitraum in etwa der heutigen Zusammensetzung in vielen OECD-Staaten entspricht, sieht der Autor diesen für negative Anreizwirkungen einer gemeinsamen Besteuerung sprechenden Befund auch als Anhaltspunkt für die aktuelle Debatten.

Eine Allgemeingültigkeit dieser Ergebnisse und damit deren Übertragung auf die aktuelle Situation in Deutschland darf jedoch angezweifelt werden, da die gemessenen Reaktionen auf Steuern sowohl international als auch über die Zeit stark divergieren (vgl. Bargain/Peichl/Orsini (2014)). So wurde in der Arbeit von Blau/Kahn (2007) für die USA eine über die Zeit deutlich sinkende Arbeitsangebotselastizität bei verheirateten Frauen festgestellt, wobei Kumar/Liang (2016) diesen Befund lediglich im Hinblick auf die Partizipation am Arbeitsmarkt bestätigen.

Zuletzt ist festzustellen, dass dem (langfristig angelegten) Eheversprechen tatsächlich eine grundsätzliche Aufgabenteilung innerhalb des Haushaltes zu folgen scheint, was bei der Interpretation geringerer Marktarbeitsstunden nach der Heirat und des kausalen Zusammenhanges zum Ehegattensplitting berücksichtigt werden sollte. So zeigen Baumgarten/Houben (2014) zwar, dass Frauen und Männer in der Ehe etwa gleiche Teile zur Arbeit beitragen, allerdings verwenden Ehefrauen im Verhältnis mehr Zeit auf die Hausarbeit sowie ggf. die Kinderbetreuung, während Ehemänner mehr Zeit auf Erwerbsarbeit verwenden. Barg/Beblo (2012) kommen zu ähnlichen Erkenntnissen hinsichtlich der Aufteilung in der Ehe, wobei Frauen insgesamt sogar mehr Zeit für Erwerbs- und Hausarbeit sowie Kinderbetreuung investieren als Männer.³⁰ Gleichzeitig liegt die Erwerbsbeteiligung verheirateter Frauen in Deutschland unterhalb der von alleinstehenden Frauen.³¹

3 Das Ansatz dieser Arbeit

Da es in den letzten Jahrzehnten keine zum Revenue Act 1948 vergleichbare Steuerreform in Deutschland gab, die einen Teil der Ehepaare betraf andere dagegen nicht, soll hier eine alternative Identifikationsstrategie gewählt werden: Bei einer Anpassung des Arbeitsverhaltens aufgrund des Ehegattensplittings sollte nach der Heirat, welche hier als Treatment dient, bei Frauen eine signifikante Reduktion der für den Beruf aufgewendeten Stunden, die in dieser Analyse die Ergebnisgröße darstellen, zu beobachten sein.

Der dabei beobachtete Effekt wird den tatsächlichen Effekt des Ehegattensplittings tendenziell überschätzen. Eine Aufgabenteilung bzw. gemeinsame Planung der Erwerbs- und Hausarbeit nach der Eheschließung oder eine allgemeine Reduzierung der Erwerbstätigkeit des Zweitverdieners dürfte neben dem Ehegattensplitting durch erhebliche Absicherungseffekte, die mit der Eheschließung einhergehen (höhere Kosten der Beendigung der Partnerschaft; den

³⁰ Außerdem zeigen Sie, dass die Aufteilung bei zusammenlebenden Paaren zwar ausgeglichener ist, die Männer hier allerdings trotzdem einen höheren Anteil ihrer Zeit mit Erwerbsarbeit verbringen.

³¹ Nichtverheiratete Frauen: 76,7%, verheiratete Frauen: 70,1%; siehe Maiterth/Chirvi (2015), Fn. 152. In dem Report von Cornelißen (2005), S. 120, wird mit Daten aus 2004 gezeigt, dass in Ostdeutschland die Erwerbstätigenquote verheirateter Frauen sogar oberhalb der der ledigen liegt.

Zugewinnausgleich und Versorgungsausgleich im Scheidungsfall), sowie sozialversicherungsrechtliche Regelungen (Familienversicherung und Witwenrente) beeinflusst werden. Die Ergebnisse der vorliegenden Analyse bilden daher eine Obergrenze für einen möglichen Effekt des Ehegattensplittings.

Ein einfacher Vorher-Nachher-Vergleich³² des Arbeitsverhaltens von Frauen (Zweitverdienern) vor und nach der Heirat³³ birgt u. a. das Risiko, dass andere Einflussfaktoren im Untersuchungszeitraum nicht ohne weiteres von dem Effekt der Heirat getrennt werden können. Neben dem generellen über die vergangenen Jahrzehnte kontinuierlichen Anstieg der Erwerbsbeteiligung von Frauen könnten auch konjunkturelle Einflüsse, die auf das allgemeine Beschäftigungsniveau wirken, verzerrend wirken. Das Arbeitsverhalten mag insbesondere um das typische Heirats- und Mutterschaftsalter von etwa 30 Jahren zudem einer grundsätzlichen, individuellen Veränderung unterliegen. Um derartige Einflussfaktoren auszuschließen, wird der Heiratseffekt daher mittels eines Differenz-in-Differenzen³⁴-Ansatzes unter Berücksichtigung einer geeigneten Kontrollgruppe identifiziert. Die eheähnlichste Partnerschaftsform und daher die Basis für die Kontrollgruppe ist die nichteheliche Lebensgemeinschaft, kurz NEL.³⁵

Die Kontrollgruppe bilden hier diejenigen Frauen, die im Zeitraum, in dem die Frauen aus der Treatmentgruppe heiraten, mit einem festen Partner in einem gemeinsamen Haushalt leben. Es wird daher angenommen, dass sich das Arbeitsverhalten der heiratenden Frauen ohne die Eheschließung analog zu dem von vergleichbaren, in NEL lebenden Frauen entwickelt hätte. Als Datengrundlage dient hierzu das Sozioökonomische Panel (SOEP), in dem neben Erwerbs- und Einkommensangaben auch Partnerschaftsverläufe sowie das Führen eines gemeinsamen Haushalts beobachtet werden können. Neben der Identifikation einer hinreichenden Ähnlichkeit der beiden Gruppen, zu der weitere Ausführungen in einer umfangreicheren Version dieses Aufsatzes zu finden sind, und einer weiter unten vorzufindenden genaueren Beschreibung der verwendeten Daten soll an dieser Stelle auf zwei diskussionswürdige Punkte im Zusammenhang mit dem in dieser Analyse verwendeten Treatment, der Heirat, hingewiesen werden.

Um die Breite der Datenbasis für diese Analyse zu verbessern – die Anzahl derjenigen Frauen im SOEP, die pro Jahr heiraten und der Treatmentgruppe zugeordnet werden können, liegt jeweils im zweistelligen Bereich – werden in verschiedenen Jahren heiratende Paare betrachtet werden.

³² Bezogen auf die Eheschließung.

³³ Als Ergebnisgröße im kontrafaktischen Zustand „ohne Heirat“ wird hier einfach der Erwerbsstatus im Zeitpunkt vor der Heirat genommen.

³⁴ Ab hier DiD.

³⁵ Eheähnliche partnerschaftliche Lebensformen werden auch als „quasi marriages“ bezeichnet, vgl. Adamopoulou (2010). Eine umfassendere Darstellung der NEL sowie der Unterschiede und Ähnlichkeiten zu Ehen finden sich in einer umfangreicheren Version dieses Aufsatzes, der beim Autor angefragt werden kann. Für einen auch dort vernachlässigten rechtlichen Überblick zu NEL siehe Grziwotz (2014).

Es stellt sich die Frage, ob als Kontrollgruppe nur diejenigen Frauen dienen sollen, die niemals heiraten werden³⁶, oder ob eine spätere Heirat und daher potentielle Zugehörigkeit zur Treatmentgruppe zu einem späteren Zeitpunkt unschädlich ist. Abgesehen davon, dass generell für eine Heirat offene Frauen die ähnlichere Kontrollgruppe bilden mögen und bei Nichtbeachtung später heiratender Frauen eine Verzerrung dieser Gruppe in Richtung einer geringeren Heiratswahrscheinlichkeit erfolgen würde, ist dieses Thema im Zusammenhang anderer Forschungsfragen bereits von Sianesi (2004), Frederiksson/Johansson (2008) oder Fitzenberger/Sommerfeld/Steffes (2013) diskutiert worden. Da ein Ausschluss später heiratender Frauen zu verzerrten Schätzungen aufgrund einer von zum jeweiligen Zeitpunkt „unbeobachtbare Eigenschaften“ geleiteten Selektion führen würde, sollten sämtliche jeweils in NEL lebende Frauen Teil der Kontrollgruppe bleiben. Dieses Vorgehen wurde auch für diese Analyse gewählt.

Ein zweites Thema ist die für die Identifikation des Heiratseffekts notwendige Annahme der bedingten Unabhängigkeit³⁷ der (Entscheidung zur) Heirat. Es darf also keine bei der Schätzung unberücksichtigt bleibenden Einflussfaktoren geben, aufgrund derer sich Frauen für eine Heirat entscheiden, die auch auf das Arbeitsverhalten wirken. Da es sich bei der Heirat um kein zufälliges Ereignis handelt – wie im Idealfall eines randomisierten Experiments – verdient dieser Punkt besondere Aufmerksamkeit. Fitzenberger/Sommerfeld/Steffes (2013) wählen in ihrer Arbeit beispielsweise die Geburt eines Kindes als Treatment und argumentieren, bei Berücksichtigung reichhaltiger Kontrollvariablen, die wie in der vorliegenden Arbeit auch Persönlichkeitsmerkmale enthalten, sei die Annahme einer bedingten Unabhängigkeit des Treatments erfüllt. Folgt man dieser Argumentation, lässt sich auch das Problem der Selbstselektion in die Ehe durch umfangreiche Kontrollvariablen lösen und die Annahme der bedingten Unabhängigkeit kann als erfüllt angesehen werden. Sollten dennoch unberücksichtigte, den Heiratswunsch beeinflussende Faktoren vorliegen, dürften dies in erster Linie solche sein, die das Arbeitsverhalten der heiratenden Frauen tendenziell negativ beeinflussen. Dies trifft für eine möglicherweise existierende Präferenz heiratender Frauen für eine traditionelle Arbeitsteilung ebenso zu wie für einen stärkeren Wunsch nach finanzieller Absicherung. Letztere ist bei verheirateten Frauen durch die zivilrechtlichen Regelungen im Zusammenhang mit dem Institut „Ehe“ (Zugewinn- und Versorgungsausgleich) ungleich stärker ausgeprägt als bei nicht-verheirateten Frauen. Eine existierende Endogenität der Heirat, die durch die Kontrollen nicht aufgefangen wird, wird die hierdurch auftretende Verzerrung der Schätzergebnisse den wahren Effekt tendenziell überschätzen.

³⁶ Dieses Merkmal könnte in aktuellen Jahren nicht festgestellt werden.

³⁷ Die sogenannte „conditional independence assumption“, kurz CIA.

Die Tatsache, dass mit der Eheschließung häufig die Geburt eines Kindes einhergeht, spielt im Rahmen der Identifikationsproblematik eine untergeordnete Rolle, da dieses Merkmal beobachtbar ist. Problematisch wäre es nur, falls der Kinderwunsch durch die Möglichkeit des Ehegattensplittings veranlasst wäre, wovon jedoch nicht auszugehen ist. Daher nehme ich an, dass die Entscheidung, ein Kind zu bekommen, exogen und nicht durch das Ehegattensplitting verursacht ist.

Obwohl eine ersatzlose Streichung des Ehegattensplittings in Deutschland verfassungsrechtlich schwer möglich scheint³⁸, soll diese Arbeit grundsätzlich analysieren, ob sich das Ehegattensplitting im Vergleich zur Individualbesteuerung signifikant negativ auf das Arbeitsangebot verheirateter Frauen auswirkt. Das Ergebnis wird daher eine Obergrenze zu erwartender Auswirkungen einer wie auch immer gestalteten Begrenzung des Ehegattensplittings darstellen.

4 Daten

Die vorliegende Arbeit basiert auf Daten des Sozioökonomischen Panels, einer im Jahr 1984 begonnenen jährlichen repräsentativen Panelumfrage unter privaten Haushalten in Deutschland.³⁹ Für die vorliegende Analyse wurden die Befragungswellen 2002 bis 2012 verwendet, auf Daten weiter zurückliegender Jahre wurde im Hinblick auf eine möglichst große Validität für aktuell diskutierte Reformoptionen verzichtet. Wie bereits beschrieben wurde, unterliegt insbesondere die Arbeitsangebotsentwicklung von Frauen einem stetigen Wandel, dennoch scheint es mit Blick auf eine möglichst umfangreiche Datengrundlage vertretbar, bis 2002 zurückreichende Daten zu verwenden.

Als Untersuchungszeitraum für die DiD-Analyse wurden fünf aufeinanderfolgende Jahre gewählt. Es wird die Veränderung der für den Beruf aufgewendeten Stunden⁴⁰ vom ersten (vor der potentiellen Heirat; t_{vor}) bis zum fünften Jahr (nach der potentiellen Heirat; t_{nach}) untersucht, während im dritten Jahr das potentielle Treatment, die Heirat, stattfindet. Gemäß der hier

³⁸ Vgl. Maiterth/Chirvi (2015), S. 21. Das BVerfG fordert so bspw. in einem Urteil vom 07.05.2013 die Steuerfreistellung des familiären Existenzminimums. Gleichzeitig würde ein Splitting mit einem übertragbaren Betrag unterhalb des sog. Realsplittings im derzeitigen Recht verheiratete Paare gegenüber geschiedenen Paaren schlechterstellen, so dass dies ohne Änderung des Realsplittings die derzeit minimal zulässige Splittingform darstellen mag.

³⁹ Genauere Informationen zum SOEP finden sich z. B. bei Wagner/ Fricke/ Schupp (2007), Wagner/Göbel/Krause/Pischner/Sieber (2008) oder allgemein unter <http://www.diw.de/de/soep>.

⁴⁰ Im SOEP sind hierzu die Stunden pro Tag angegeben. Für die Auswertungen wurden diese Beträge mit fünf multipliziert, um Angaben zu häufig ausgewiesenen und besser einzuschätzenden Wochenstunden zu erhalten.

getroffenen Annahme paralleler Trends, die dem DiD-Ansatz zu Grunde liegt, würde der Trend des Arbeitsangebots der im dritten Jahr heiratenden (Treatment-)Gruppe ohne die Heirat so verlaufen wie jener der Kontrollgruppe. Jegliche Abweichung von diesem Trend wird, nach Kontrolle auf Unterschiede zwischen den Gruppen (dazu weiteres im folgenden Kapitel), der Heirat zugeschrieben. Da, wie bereits ausgeführt, neben dem Ehegattensplitting noch andere Effekte der Heirat negativ auf das Arbeitsverhalten von Ehefrauen wirken sollten, kann zumindest bei einem unveränderten Arbeitsangebot verheirateter Frauen ein Einfluss des Ehegattensplittings ausgeschlossen werden. Im umgekehrten Fall ist ein negativer Effekt des Ehegattensplittings zumindest nicht auszuschließen.

Durch den für die Analyse gewählten Zeitraum ist einerseits sichergestellt, dass auch verzögerte Änderungen des Arbeitsangebotes gemessen werden können (zwischen dem Heirats- und dem zweiten Beobachtungszeitpunkt liegt noch ein volles Jahr). Andererseits ist der Zeitraum kurz genug, um die Gefahr der Verzerrung der Ergebnisse durch andere (exogene) Einflüsse zu minimieren, die bei einem längeren Betrachtungszeitraum auftreten könnte. Er stellt insofern einen Kompromiss zwischen einer ausreichenden Zeit zur Messung von Verhaltensanpassungen sowie einer möglichst isolierten Identifizierung des Heiratseffekts von sonstigen Einflussfaktoren dar. Außerdem dürfte die Heirat in aller Regel keine spontane Entscheidung sein, sondern einer mehr oder weniger langfristigen Planung folgen. Entgegen vieler anderer Quasi-Experimente gibt es einen Effekt hier also möglicherweise bereits direkt im Jahr der Heirat bzw. sogar davor.⁴¹ Eine das Arbeitsverhalten beeinflussende Wirkung der Eheschließung sollte also zeitnah zu beobachten sein.

Die SOEP-Population wurde für Zwecke der vorliegenden Untersuchungen begrenzt. Es werden lediglich diejenigen Frauen berücksichtigt, die entweder in den gesamten fünf Jahren mit demselben Partner zusammengelebt haben (Kontrollgruppe) oder für die dies in den ersten beiden Jahren der Fall war und die – nach Eheschließung im dritten Jahr – in den Jahren vier und fünf mit diesem Partner verheiratet sind (Treatmentgruppe). Weiter werden nur Frauen betrachtet, die sowohl im Jahr eins als auch im Jahr fünf Angaben zu der Anzahl der Stunden gemacht haben, die sie täglich für ihren Beruf aufwenden. Außerdem werden nur Frauen berücksichtigt, die im Jahr fünf mindestens 25 und höchstens 60 Jahre alt sind, um Verzerrungen durch Ein- bzw. Austritt in bzw. aus dem Arbeitsmarkt zu vermeiden. Bei Betrachtung der Abbildungen in Abschnitt 6 sollte bedacht werden, dass Daten zu den Jahren außerhalb des genannten 5-Jahres-Zeitraum von t_{vor} bis t_{nach} nicht bei allen Beobachtungen vorliegen.

⁴¹ Dies sollte natürlich bei der Überlegung zum Messzeitpunkt vor der Heirat berücksichtigt werden. Die folgende deskriptive Analyse zeigt, dass im gewählten Jahr t_{vor} noch keine Wirkungen festzustellen sind, im Jahr vor der Heirat allerdings schon.

t_{vor}	Heiratsjahr (Treatment)	t_{nach}	Anzahl Beobachtungen	
<i>Jahr</i>	<i>Jahr</i>	<i>Jahr</i>	<i>Treat</i>	<i>Kontroll</i>
2002	2004	2006	52	182
2003	2005	2007	70	201
2004	2006	2008	73	246
2005	2007	2009	85	246
2006	2008	2010	77	237
2007	2009	2011	59	191
2008	2010	2012	59	207
			475	1510

Abbildung 1: Anzahl der Beobachtungen in den einzelnen Beobachtungszeiträumen.

Es ergeben sich die in Abbildung 1 dargestellten sieben „DiD-Datensätze“ für den Gesamtzeitraum 2002 bis 2012. Bei einer Untersuchung der einzelnen Datensätze ist die Zahl der Beobachtungen gerade in den Treatment-Gruppen relativ gering. Durch ein Matching innerhalb der einzelnen Datensätze können die Daten allerdings gemeinsam betrachtet werden, ohne dass der zeitliche Bezug der Beobachtungen beider Gruppen verloren geht. Die monetären Größen (Einkommen, Lohn, Vermögen) wurden für diese gemeinsame Untersuchung inflationsbereinigt. Aufgrund auch in den Beobachtungsjahren schon bestehender Mindestlöhne in vielen Bereichen und teilweise unrealistisch niedriger Stundenlöhne, wurden Werte unter drei Euro pro Stunde gelöscht und neu geschätzt. Hierzu folgen weitere Erläuterungen weiter unten.

5 Deskriptive Auswertung

Im Folgenden werden zunächst alle erhobenen Daten gemeinsam ausgewertet, um die Standardfehler aufgrund kleiner Beobachtungszahlen nicht „künstlich“ zu vergrößern. Die deskriptive Auswertung enthält aufgrund der Berücksichtigung aller Datensätze Beobachtungen unter Umständen mehrfach. In einem zeitlich späteren Datensatz beobachtete heiratende Frauen können sich bspw. in zeitlich früheren Datensätzen in der Kontrollgruppe befinden.

Betrachtet man die von den Frauen beider Gruppen für den Beruf aufgewendeten Stunden im Zeitraum von t_{vor} bis t_{nach} , ergibt sich der Verlauf aus Abbildung 2. Zusätzlich werden an dieser Stelle zur besseren Einschätzung der Entwicklung die Berufsstunden jeweils drei Jahre vor und

nach dem analysierten Zeitraum dargestellt.⁴² Es ist deutlich zu erkennen, dass die Verläufe der Berufsstunden in den Jahren vor der (potentiellen) Heirat in beiden Gruppen sehr ähnlich sind. Dies ist eine wichtige Erkenntnis hinsichtlich der für DiD-Analysen zentralen Annahme paralleler Trends.⁴³ Heiratende Frauen reduzieren ihre Berufsstunden nach der Eheschließung allerdings deutlich, während diese bei in NEL lebenden Frauen relativ stabil bleiben. Daher scheint die Eheschließung und damit verbunden das Ehegattensplitting einen negativen Einfluss auf die Erwerbstätigkeit von Frauen zu auszuüben. Um zu identifizieren, ob dies auch tatsächlich der Fall ist, muss der Heiratseffekt von anderen Einflüssen isoliert werden. Dazu müssen anderweitige Unterschiede zwischen den Gruppen identifiziert und idealiter vollständig isoliert werden. Im nächsten Schritt erfolgt daher eine deskriptive Analyse der Charakteristika beider Gruppen, die sich an den Erkenntnissen der bisherigen Literatur orientiert.

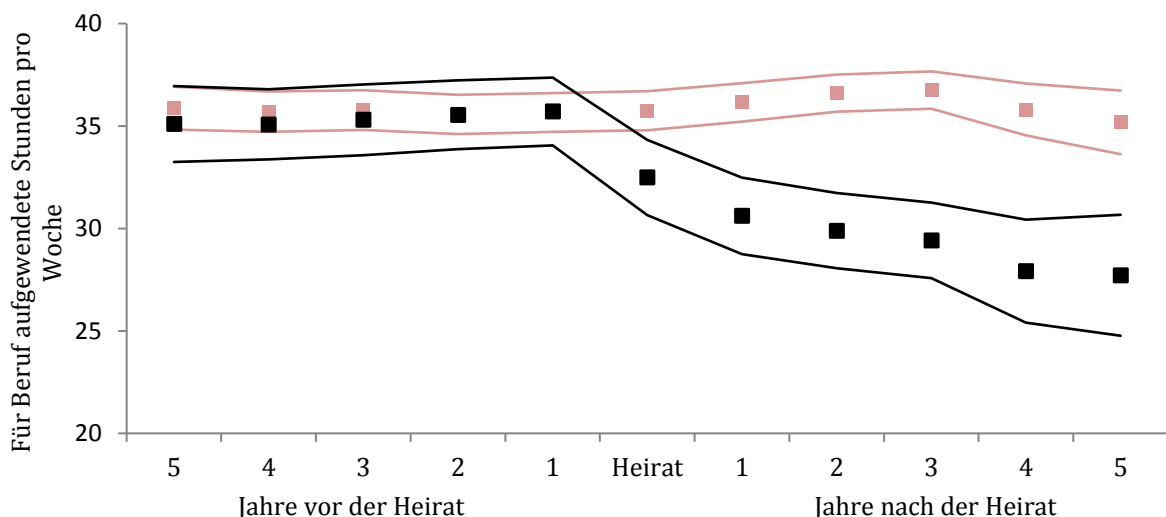


Abbildung 2: Für den Beruf aufgewendete Wochenstunden der heiratenden Frauen (schwarz) im Zeitverlauf. In hellrot dazu im Vergleich die Wochenstunden der Kontrollgruppe im identischen Zeitraum; dazu 95%-Konfidenzintervall. Quelle: Eigene Berechnungen auf Basis des Datensatzes „Berufstätigkeit“; SOEP-Wellen 2002 – 2012.

Naturgemäß lassen sich nicht sämtliche Charakteristika, in denen sich die Gruppen unterscheiden könnten, beobachten (z. B. (beruflicher) Ehrgeiz, Konservativismus) oder sie liegen in den SOEP-Daten nicht vor. Auch wenn man versuchen kann, einzelne unbeobachtete Eigenschaften durch

⁴² Da es für die Berücksichtigung im Datensatz nur Voraussetzung ist, im fünfjährigen Beobachtungszeitraum Angaben zur Partnerschaftsform und den Berufsstunden gemacht zu haben, sind die Angaben zu Berufsstunden vor und nach diesem Zeitraum nicht voll besetzt.

⁴³ Engl. „parallel trend assumption“. Ein Approximieren des kontrafaktischen Verlaufs der Berufsstunden der Treatmentgruppe anhand des Verlaufs der Kontrollgruppe und somit eine Schätzung des Effekts der Heirat ist grundsätzlich nur möglich, wenn die Berufsstunden in beiden Gruppen bei Abwesenheit der Heirat gleich verlaufen würden.

vorhandene Merkmale annäherungsweise abzubilden (z. B. Konservatismus durch das Wahlverhalten oder die Religionszugehörigkeit), ist eine Berücksichtigung sämtlicher unbeobachtbarer Einflussfaktoren ausgeschlossen. Unberücksichtigte, die Ergebnisgröße (unterschiedlich) beeinflussende Eigenschaften finden sich daher im Fehlerterm einer Regression wieder und verzerren die Parameter anderer Variablen (inkl. des DiD-Schätzers). Eine Herausforderung besteht daher darin, auch den Einfluss von Unterschieden nicht beobachtbarer Einflussfaktoren zu erfassen, um zu validen Regressionsergebnissen zu gelangen.

Einen Lösungsansatz hierzu bietet die psychologische Persönlichkeitsforschung. Diese hat eine „weithin anerkannte Taxonomie der Persönlichkeit“⁴⁴ hervorgebracht. Die Gesamtheit der Persönlichkeitsmerkmale (personality traits) wird hier in fünf Dimensionen dargestellt (Neurotizismus, Extraversion, Offenheit gegenüber neuen Erfahrungen, Gewissenhaftigkeit und Verträglichkeit⁴⁵) mit dem Ziel, ein sparsames Beschreibungssystem sämtlicher Persönlichkeitsmerkmale zu erhalten. Die Ausgangsgröße dieses Systems besteht aus ehemals knapp 18.000 persönlichkeitsbeschreibenden Worten, lexikalisch ausgewählt von Allport/Odbert (1936), die Reduktion auf die fünf genannten Dimensionen wurde indes in verschiedenen Arbeiten auf unterschiedlichen Wegen erreicht und „bestätigt“.⁴⁶

Der Umfang der „Befragungssitems“ zur Persönlichkeit wurde für SOEP-Zwecke auf 15 (ab der zweiten Befragungswelle 16) von Gerlitz/Schupp (2005) entwickelte Items reduziert („BFI-S“)⁴⁷, da die Abfragezeit des für den regelmäßig in der psychologischen Forschung verwendeten „NEO-PI-R“-Katalog mit 240 Einzelfragen mit dem SOEP-Umfrageformat nicht kompatibel gewesen wäre.⁴⁸ Aufgrund der nicht jährlichen Erfassung der den fünf Dimensionen zu Grunde liegenden Befragungspunkte – die Erfassung erfolgte bisher mit den Befragungswellen 2005, 2009 sowie 2013 – sowie eines geänderten Designs ab 2009 werden in der vorliegenden Arbeit einheitlich die Ergebnisse der BIG5-Befragung aus dem Jahr 2009 verwendet.

Eine Zusammenfassung der Charakteristika beider Gruppen findet sich in Abbildung 3. Dabei ist festzustellen, dass beide Gruppen im Gesamtblick durchaus vergleichbar sind. Heiratende Frauen sind lediglich signifikant⁴⁹ religiöser, ihr Anteil ist höher in den alten Bundesländern, sie haben weniger Berufserfahrung, arbeiten seltener in Führungspositionen und sind jünger, wobei sich

⁴⁴ So z. B. Dehne/Schupp (2007), S. 1.

⁴⁵ Siehe z.B. Asendorph/Neyer (2012), S. 107.

⁴⁶ Siehe ebd., S. 107 - 108. Tlw. wurden diese Dimensionen ergänzt oder reduziert. Vgl. auch Borkenau/Ostendorf (2008).

⁴⁷ Eine Validierung dieses Konzepts nehmen auch Dehne/Schupp (2007) vor. Eine aktuellere Übersicht gibt es bei Schupp/Gerlitz (2014).

⁴⁸ Nach Dehne/Schupp (2007), S. 31 - 32.

⁴⁹ Mit „signifikant“ ist im Fließtext dieser Arbeit, soweit nicht anders ersichtlich, jeweils Signifikanz auf 5%-Niveau gemeint.

die Bedeutung des letztgenannten Befundes bei einer weiter unten folgenden genaueren Betrachtung relativiert.

		Treatmentgruppe	Kontrollgruppe	t	Pr
<u>Allgemein</u>					
Alter	<i>Jahre</i>	33,60	37,54	8,27	0,000
Alter (nur > 35 Jahre)	<i>Jahre</i>	43,26	44,05	1,59	0,111
Alter (nur ≤ 35 Jahre)	<i>Jahre</i>	28,21	28,47	1,01	0,311
Altersdifferenz zum Partner	<i>Jahre</i>	2,78	2,54	-0,74	0,457
Anteil religiös	<i>Prozent</i>	63,16	51,52	-4,46	0,000
Bildungsniveau	<i>1 - 4</i>	3,16	3,17	0,12	0,908
Neue Bundesländer	<i>Prozent</i>	33,68	39,87	2,42	0,016
Städtischer Wohnraum	<i>Prozent</i>	62,95	67,75	1,94	0,053
<u>Kinder</u>					
Kinder	<i>Anzahl</i>	0,41	0,43	0,57	0,566
Kind geboren im Beobachtungszeitraum	<i>Prozent</i>	38,11	14,24	-11,73	0,000
<u>Arbeit, Einkommen und Vermögen</u>					
Beruf Partner	<i>Std/Woche</i>	42,46	39,66	-2,40	0,016
Anteil Einkommen	<i>Prozent</i>	40,73	41,93	0,87	0,385
Beruf	<i>Std/Woche</i>	35,55	35,57	0,02	0,985
Berufserfahrung (TZ+VZ)	<i>Jahre</i>	10,58	14,02	7,30	0,000
Betrieb > 200 MA	<i>Prozent</i>	39,28	37,48	-0,64	0,525
Führungsposition	<i>Prozent</i>	13,05	17,25	2,17	0,031
Hausarbeit	<i>Std/Woche</i>	9,14	8,94	-0,59	0,558
Hausarbeit Partner	<i>Std/Woche</i>	4,00	4,10	0,38	0,704
Haushaltseinkommen	<i>Euro/Monat</i>	2.794	2.893	1,15	0,248
Öffentlicher Dienst	<i>Prozent</i>	31,20	30,00	-0,45	0,652
Stundenlohn	<i>Euro</i>	14,27	14,37	0,25	0,806
Vermögen	<i>Euro</i>	53.999	50.789	-0,42	0,677
Vermögen Partner	<i>Euro</i>	65.823	80.114	1,26	0,210
Zufrieden im Job	<i>1 - 11</i>	7,77	7,79	0,17	0,866
<u>Persönlichkeitsmerkmale:</u>					
Extraversion	<i>1 - 7</i>	4,88	4,96	1,22	0,221
Gewissenhaftigkeit	<i>1 - 7</i>	5,85	5,94	1,94	0,053
Neurotizismus	<i>1 - 7</i>	4,12	3,92	-3,21	0,001
Offenheit	<i>1 - 7</i>	4,64	4,69	0,89	0,372
Verträglichkeit	<i>1 - 7</i>	5,26	5,39	2,45	0,014

Abbildung 3: Ausgewählte Charakteristika der beiden Gruppen; ungewichtet; grundsätzlich zum Zeitpunkt t_{vor} ; „Kind bekommen im Beobachtungszeitraum“ bezieht sich auf den Zeitraum t_{vor} bis t_{nach} ; die Persönlichkeitsmerkmale stammen aufgrund nicht jährlicher Befragung alle aus dem Jahr 2009. Fett gedruckte Angaben weisen auf signifikante (5%-Niveau) Unterschiede zwischen den Gruppen hin.

Auch bezogen auf die oben genannten Persönlichkeitsmerkmale unterscheiden sich die Frauen beider Gruppen zum Teil. Betrachtet man die abgefragten „Items“ genauer, finden sich eine signifikant selbstbewusstere Einschätzung der Qualität der eigenen Arbeit und der Wertschätzung von Kunst und Kultur bei der Kontrollgruppe, während heiratende Frauen sich als signifikant nachtragender einschätzen, sich größere Sorgen machen und schlechter Stress verarbeiten können. Wie in Abbildung 3 ersichtlich, unterscheiden sich die Frauen beider Gruppen jedoch nicht signifikant hinsichtlich ihrer für den Beruf verwendeten Zeit zum Zeitpunkt t_{vor} ; eine für die Untersuchung schädliche Antizipation der Heirat hat also noch nicht stattgefunden.

Betrachtet man die Verteilungen der Ausprägungen, stellt man beim Alter der heiratenden Frauen eine im Ansatz bimodale Verteilung mit Modi bei knapp unter 30 sowie etwas über 40 Jahren fest.⁵⁰ Eine ähnliche, wenn auch weniger eindeutige Verteilung findet sich bei den in NEL lebenden Frauen.⁵¹ Der Anteil der „älteren“ Frauen in der Kontrollgruppe ist im Vergleich allerdings deutlich größer, wodurch das höhere Durchschnittsalter erklärt werden kann. Bei Teilung beider Gruppen anhand einer Altersgrenze von 35 Jahren unterscheidet sich das Alter zwischen Treatment- und Kontrollgruppe weder unter- noch oberhalb dieser Grenze signifikant. Die sich signifikant unterscheidende Berufserfahrung hängt indes sehr eng mit dem Alter zusammen und ist durch den im Vergleich höheren Anteil der Frauen im Alter von mehr als 35 Jahren in der Kontrollgruppe zu erklären. Nach Unterteilung in die zwei Altersgruppen ist auch die Berufserfahrung jeweils nicht mehr signifikant unterschiedlich.⁵²

Gleichzeitig ist nach der deskriptiven Auswertung bei den über 35-jährigen Frauen auch nach der Heirat kaum eine Veränderung der Berufsstunden zu erkennen. Dies hängt unter Umständen auch mit einem anderen, nicht erst nach den Erkenntnissen aus den vorangegangenen Kapiteln wenig überraschenden Ergebnis zusammen: Während sich die Anzahl der Kinder in den beiden Vergleichsgruppen im Zeitpunkt t_{vor} nicht signifikant unterscheiden, bekommt ein deutlich größerer Teil der heiratenden Frauen (fast 40 %) im Beobachtungszeitraum ein Kind. Das durchschnittliche Alter der Frauen, die in diesem Zeitraum ein Kind bekommen liegt in beiden Gruppen bei ziemlich genau 31 Jahren bezogen auf den Zeitpunkt t_{vor} , die Geburtenrate bei Frauen im Alter von mehr als 35 Jahren ist dabei allerdings gering.

Betrachtet man lediglich die Berufsstunden derjenigen Frauen, die im Zeitraum von t_{vor} bis t_{nach} kein Kind bekommen, ergibt sich bereits deskriptiv ein deutlich anderes Bild. Wie Abbildung 4

⁵⁰ Bei den anderen Variablen unterscheiden sich die Verteilungen zwischen den beiden Gruppen nicht grundsätzlich.

⁵¹ Hier befindet sich die zweite Konzentration allerdings knapp unter 40 Jahren.

⁵² Bei den maximal 35-jährigen stimmen die Werte bis auf die erste Nachkommastelle überein, während die Berufserfahrung in der älteren Gruppe zumindest auf 5%-Signifikanzniveau nicht unterschiedlich ist.

zeigt⁵³, kommt es bei heiratenden Frauen zu einem bestenfalls minimalen Rückgang der Berufsstunden im Zuge der Heirat. Die in Abbildung 2 eindeutig vorhandene Reduktion der Arbeit scheint also vor allem durch Kinder ausgelöst zu sein, wenngleich Folgen der Heirat – z. B. das Ehegattensplitting – diese verstärken mögen.

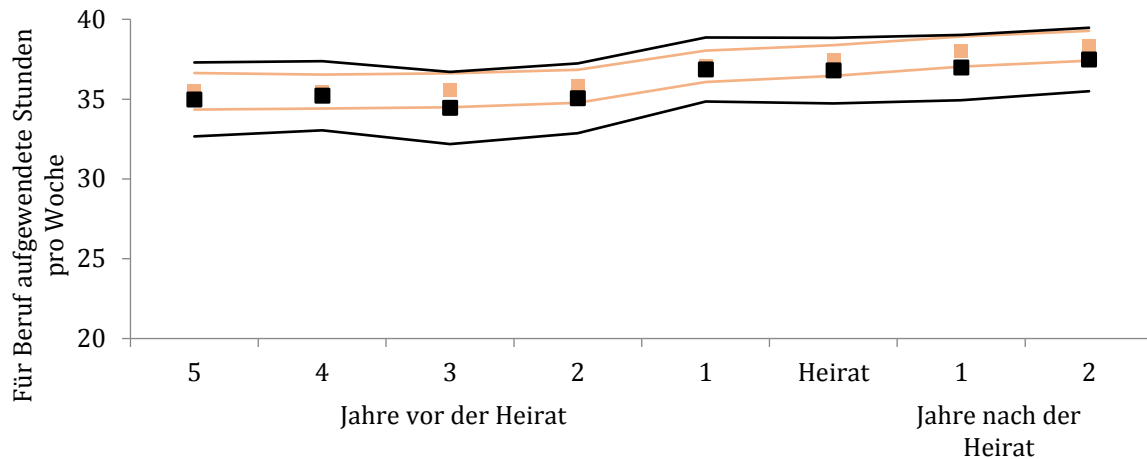


Abbildung 4: Für den Beruf aufgewendete Wochenstunden im Zeitverlauf mit Begrenzung auf die Frauen, die im Zeitraum t_{vor} bis t_{nach} kein(e) Kind(er) bekommen haben. Heiratende Frauen (Treatmentgruppe) in schwarz. Zum Vergleich die Werte der Kontrollgruppe in hellrot; dazu 95%-Konfidenzintervall. Quelle: Eigene Berechnungen auf Basis des Datensatzes „Berufstätigkeit“; SOEP-Wellen 2002 – 2012.

Abbildung 5 gibt dazu Aufschluss über den Anteil der Frauen beider Gruppen, die in den Jahren um die Heirat Kinder bekommen haben. Dabei wird deutlich, und dies stützt die bisherigen Ausführungen zur Bedeutung von Kindern, dass die heiratenden Frauen vermehrt um den Zeitraum der Heirat ein Kind bekommen, während es bei Frauen in der Kontrollgruppe wenig überraschend eine gleichmäßigere Verteilung über die betrachteten Jahre gibt.

Gleichzeitig führt die Tatsache, dass Ehepaare auch im Zeitpunkt t_{nach} und danach deutlich häufiger Kinder bekommen als in NEL lebende Paare, dazu, dass die rein ehebedingte Reduktion der Berufsstunden im Weiteren überschätzt wird, sofern nicht auch auf das Merkmal eines nach t_{nach} geborenen Kindes kontrolliert wird.⁵⁴

⁵³ Für diese Abbildung wurden die Jahre nach t_{nach} nicht mehr betrachtet, da neben der Abnahme der Anzahl der Beobachtungen der Verlauf durch Frauen verzerrt wird, die in den Folgejahren Kinder bekommen – auch dies ist deutlich häufiger bei heiratenden Frauen der Fall.

⁵⁴ Die Geburt eines Kindes nach t_{nach} wird den im Vergleich zur Kontrollgruppe leicht abfallenden Verlauf in Abbildung 2 beeinflusst haben, da hier nur Frauen aus der Stichprobe genommen wurden, die ein Kind im Zeitraum t_{vor} bis t_{nach} bekommen haben, die Nachwuchsquoten bei den heiratenden Frauen aber auch in den folgenden Jahren höher liegen.

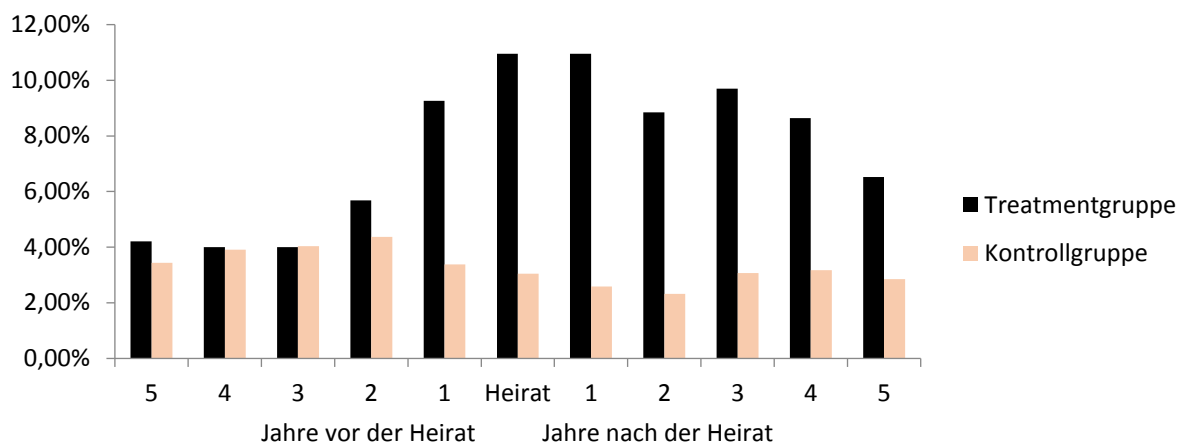


Abbildung 5: Anteil der Frauen beider Gruppen, die im jeweiligen Jahr ein Kind bekommen haben. Quelle: Eigene Berechnungen auf Basis des Datensatzes „Berufstätigkeit“; SOEP-Wellen 2002 – 2012.

6 Methodik

Wie in den Ausführungen der Langversion dieses Aufsatzes dargestellt und durch die deskriptive Analyse bestätigt, ist die Gruppe der Paare in NEL in ihren Merkmalen durchaus mit verheirateten Paaren vergleichbar. Bestehende Unterschiede müssen allerdings im Rahmen der Regressionsanalyse berücksichtigt werden, da ein naiver Mittelwertvergleich der Berufsstunden bei nicht absolut identischen Charakteristika zu falschen Ergebnissen führt.⁵⁵ Wie weiter oben ausgeführt wurde, ist die Annahme der bedingten Unabhängigkeit der Heirat bei Kontrolle auf bestehende Unterschiede überzeugend begründbar und ein mögliches Endogenitätsproblem der Heirat daher vernachlässigbar. Soweit ein Endogenitätsproblem bestehen bleibt, sprechen theoretische Überlegungen für eine Überschätzung des tatsächlichen Heiratseffektes auf das Arbeitsverhalten durch den gewählten Ansatz.

Die Regressionsanalyse basiert auf dem Zusammenhang zwischen der gesamten in die Berufstätigkeit investierten Zeit hb_{it} und den bereits in der deskriptiven Analyse beschriebenen Charakteristika, die einen Einfluss auf die Ergebnisgröße (hb_{it}) haben dürften, eines Zeit- und eines Gruppendummies sowie der DiD-Interaktionsvariable.

⁵⁵ In einem DiD-Setting ist die Berücksichtigung von Variablen als eine Kontrolle auf die Unterschiede der beiden untersuchten Gruppen zu verstehen. Es ist daher nicht notwendig auf sämtliche, die Ergebnisgrößen beeinflussenden Variablen zu kontrollieren, sondern nur auf diejenigen, in denen sich die Gruppen unterscheiden. Im Falle einer Randomisierung der Gruppen würde es so auch genügen, die Ergebnisgröße lediglich auf den Zeit- bzw. Gruppendummy sowie die DiD-Variable zu regressieren. Würde die Untersuchung anhand eines Vorher-Nachher-Vergleichs der Treatmentgruppe unternommen werden, müssten dagegen sämtliche Einflussvariablen berücksichtigt werden, da unberücksichtigte Faktoren im Fehlerterm aufgefangen würden und der Heiratsparameter – hier anhand des Zeitdummies ermittelt – verzerrt würde.

Die Regressionsgleichung lautet:

$$hb_{it} = \beta_0 + \beta_1 X_{it} + \gamma_1 zeit_t + \gamma_2 heirat_i + \delta(zeit_t heirat_i) + \varepsilon_{it}$$

mit $t = t_{vor}; t_{nach}$ bzw. $i = 1, \dots, n$. Der Regressant hb_{it} ist die oben beschriebene Ergebnisgröße, jeweils von Individuum i im Zeitpunkt t . Die Dummyvariablen $zeit$ (0 für t_{vor} ; 1 für t_{nach}) und $heirat$ (0, falls Teil der Kontrollgruppe; 1, falls Teil der Treatmentgruppe) sowie die Interaktion dieser beiden, $zeit_t heirat_i$, sind das Grundgerüst des DiD-Ansatzes - δ ist der uns interessierende DiD-Indikator. Außerdem werden im Vektor X_{it} die bereits in der deskriptiven Analyse aufgeführten Charakteristika berücksichtigt. Außer Acht gelassen wurde jedoch aufgrund häufig fehlender Beobachtungen die Altersdifferenz zwischen den Partnern. Zusätzlich aufgenommen wurde das quadrierte Alter.⁵⁶ Anstatt des gesamten Haushaltseinkommens wurde aufgrund einer möglichen Endogenität nur das nicht durch die Frau erwirtschaftete Haushaltseinkommen, welches wie der Stundenlohn aufgrund der sich hieraus ergebenden besseren Modellgüte logarithmiert in das Modell eingeht, berücksichtigt. Außerdem wurde das Merkmal „Kind geboren im Untersuchungszeitraum“ differenziert und hieraus zwei genauere Merkmale gebildet: ob ein Kind unter 3 Jahren bzw. unter 6 Jahren im Haushalt lebt. Zur Kontrolle auf Unterschiede bei der Kinderbetreuung wurde jeweils noch eine Variable der regionalen Kinderbetreuungsquote für beide Altersklassen sowie eine Interaktionsvariable mit dem jeweiligen Dummy, ob ein Kind im jeweiligen Alter vorhanden ist, eingefügt. Zuletzt wurden anstatt der in der deskriptiven Auswertung aufgeführten fünf zusammengefassten Persönlichkeitsebenen die Ausprägungen sämtlicher im SOEP abgefragten Persönlichkeitsitems eingefügt, um eine bestmögliche Kontrolle der Persönlichkeitsstruktur zu ermöglichen. Inwiefern die Aufnahme der Persönlichkeitsstruktur zu besseren Ergebnissen führt, wird an späterer Stelle geklärt.

Ein viel diskutiertes Thema in statistischen Analysen ist der Umgang mit fehlenden Daten.⁵⁷ Problematisch sind hier häufig fehlende Werte zum Stundenlohn. Angaben hierzu fehlen zwangsläufig bei nicht arbeitenden Frauen. Die Tatsache, dass der Stundenlohn auf Basis anderer Variablen errechnet werden musste⁵⁸, mag ein Grund für vereinzelt unrealistisch niedrige Stundenlöhne sein. Daher wurden Werte unter drei Euro gelöscht und der Stundenlohn jeweils neu geschätzt.⁵⁹ Problematisch bei der Schätzung von Stundenlöhnen ist eine Endogenität zu

⁵⁶ Kempe (1996), der auf Grundlage von SOEP-Daten aus dem Jahr 1993 Einflussfaktoren auf das Arbeitsangebot verheirateter Frauen untersucht, findet tatsächlich zumeist fast lineare Zusammenhänge zwischen den untersuchten Merkmalen und dem Arbeitsangebot. Beim Alter stellt er dagegen einen quadratischen Zusammenhang fest, was auch in der vorliegenden Arbeit zu besseren Ergebnissen führt.

⁵⁷ Siehe z.B. Enders (2010), der einen umfassenden Überblick hierzu gibt.

⁵⁸ Hier wurde das im SOEP generierte monatliche Bruttoerwerbseinkommen durch die tatsächlich gearbeiteten (nicht die für den Beruf aufgewendeten Stunden, die z. B. auch die Anfahrt beinhalten) Stunden geteilt. Sofern keine Angabe zu den tatsächlichen Arbeitsstunden gemacht wurde, wurden hier die vereinbarten Arbeitsstunden bzw., falls auch dieser Wert fehlt, die Stunden berücksichtigt, die dem deutschen Mittelwert des angegebenen Erwerbsumfangs (Minijob, Teilzeit, Vollzeit) entsprechen.

⁵⁹ Ähnlich auch bei Triebe (2015c).

berücksichtigender (Stunden)löhne bezogen auf die Arbeitsstunden sowie deren Verzerrung aufgrund von Selektionsproblemen. Der erste Teil des Problems entsteht, da Verdienstmöglichkeiten und Stundenlohn gemeinsam determiniert sind, z. B. mag der Stundenlohn hierdurch positiv mit der Wochenarbeitszeit korreliert sein. Spitzenlöhne gibt es weniger bei Minijobs, sondern eher dort, wo Vollzeit gearbeitet wird.⁶⁰ Dieses Problem betrifft grundsätzlich alle Beobachtungen, nicht nur die ohne angegebene bzw. plausible Werte. Ein Selektionsproblem liegt vor, weil Stundenlöhne zwangsläufig nur bei Frauen beobachtbar sind, deren Marktlohn oberhalb des eigenen Reservationslohns liegt. In Anlehnung an das Vorgehen in anderen Arbeiten⁶¹ wurde hier ein Verfahren auf Basis des Korrektionsmodells von Heckman (1979) gewählt. Hier wird zunächst die Selektion in den Arbeitsmarkt anhand des zweistufigen Heckman-Modells geschätzt, anschließend unter Berücksichtigung der inversen Mills ratios aus der Heckman-Schätzung Löhne in Anlehnung an die Funktion von Mincer (1974) berechnet und folgend bei fehlenden oder nicht plausiblen Werten in der weiteren Analyse diese geschätzten Löhne berücksichtigt.⁶² Die bei der Schätzung berücksichtigten Variablen sind das Alter, die Arbeitserfahrung in Voll- und Teilzeit linear und quadratisch, Dummyvariablen für einen Wohnort in Ostdeutschland, eine Tätigkeit in führender Position, im öffentlichen Dienst, in einem Unternehmen mit über 200 Mitarbeitern sowie das Bildungsniveau. In der Selektionsschätzung werden zudem der Familienstand sowie die Anzahl der eigenen Kinder berücksichtigt. Die Stundenlöhne gehen aufgrund der in Folge besseren Modellgüte logarithmisch in die Regression ein. Auch die Schätzung der Stundenlöhne erfolgt für jeden weiter unten untersuchten Teildatensatz gesondert.

Zuletzt stellt sich die Frage, inwiefern die weiter oben aufgeführten sieben Datensätze aus Abbildung 1 in der Analyse gemeinsam ausgewertet werden können. Der pro Jahr nur relativ kleine Teil der vom SOEP erfassten heiratenden Frauen führt zu jeweils relativ wenigen Beobachtungen in der Treatmentgruppe. Eine naive „gepoolte“ Auswertung der Daten mehrerer Jahre sollte aufgrund der fehlenden Unabhängigkeit der Datensätze mit Vorsicht betrachtet werden, da sich Frauen, die in einem „frühen“ Datensatz beobachtet werden, auch in „späteren“ Datensätzen wiederfinden. Dabei gibt es sowohl über mehrere „Zeiträume“ in NEL lebende Frauen (Kontrollgruppe) als auch aufgrund von Heirat (vereinzelt auch Trennung) die Gruppe wechselnde Frauen.

Um dieses Problem zu lösen und in der DiD-Analyse möglichst ausbalancierte Vergleichsgruppen gegenüberzustellen, basieren die nachfolgenden Ergebnisse auf (innerhalb der einzelnen Datensätze) gematchten Beobachtungen. Aufgrund der relativ geringen Beobachtungszahl der

⁶⁰ Einen positiven Zusammenhang zwischen Arbeitszeit und Stundenlöhnen zeigt z. B. Wolf (2010).

⁶¹ Z. B. in der Arbeit von Triebe (2015c).

⁶² Siehe hierzu auch das Vorgehen von Puhani (1995).

einzelnen Datensätze scheidet dabei ein exaktes Matching aus. Eine dieses Problem lösende Abwandlung hiervon ist das von Iacus/King/Porro (2011a) beschriebene *Coarsened Exact Matching* (CEM).⁶³ Beim CEM findet ein Matching auf Basis von Kombinationen ex ante festgelegter Wertintervalle aller hierfür ausgewählten Variablen Anwendung. Während für einige Variablen ein exaktes Matching sinnvoll sein kann, findet für andere Variablen nur ein „erweitert“ exaktes Matching statt. Die Anwendung dieser Methode führt auch dazu, dass bei einem Fehlen entsprechender Kontrollbeobachtungen nicht alle Treatmentbeobachtungen gematcht werden und diese daher für die Analyse unberücksichtigt bleiben. Der geschätzte Effekt ist unter Umständen also nicht mehr der durchschnittliche Effekt der Heirat bezogen auf alle Heiratenden (sog. ATET; „average treatment effect on the treated“), sondern ein „lokaler“, auf die gematchten heiratenden Frauen bezogener Treatmenteffekt. Den einzelnen Beobachtungen werden beim CEM folgende Gewichtungen zugeordnet:

$$w_i = \begin{cases} 1, & \text{falls Teil der Treatmentgruppe} \\ \frac{m_C}{m_T} * \frac{m_T^k}{m_C^k}, & \text{falls Teil der Kontrollgruppe} \end{cases}$$

Hier sind m_C bzw. m_T die Anzahl der beim Matching insgesamt berücksichtigten Beobachtungen der Treatment- bzw. Kontrollgruppe und m_C^k bzw. m_T^k die Anzahl der Beobachtungen der Treatment- bzw. Kontrollgruppe, die genau die entsprechende Kombination von Variablenausprägungen aufweisen.

Das Matching bezieht sich in der vorliegenden Arbeit auf das Bildungsniveau (zweistufig; mindestens abgeschlossene Lehre bzw. Hochschulreife oder weniger), die Geburt eines Kindes in den Jahren $t_{vor} + 1$ bis t_{nach} , den Ost-Dummy sowie das Alter (zweistufig; Grenze bei 35 Jahren). Außerdem werden, wie bereits geschrieben, Beobachtungen nur innerhalb der einzelnen DiD-Datensätze, also jener aus identischen Zeiträumen, gematcht. Um den Nachteil, dass je nach Genauigkeit und Umfang der geforderten Übereinstimmung der Ausprägungen verschiedener Eigenschaften Beobachtungen der Treatmentgruppe und so Informationen bezogen auf das Verhalten der heiratenden Frauen verloren gehen – es wird wie erwähnt ein „lokaler“⁶⁴ Treatmenteffekt geschätzt – auszugleichen, erfolgt nach dem auf wenige Variablen angewendete CEM eine Berücksichtigung von Kontrollvariablen in der Regressionsanalyse.

Diese Einteilung anhand ausgewählter Merkmale stellt einen Kompromiss zwischen möglichst vergleichbaren Gruppen und möglichst wenigen aus der Stichprobe herausfallender Beobachtungen der Treatmentgruppe dar. Von den ursprünglich 475 heiratenden Frauen bleiben

⁶³ Eine theoretische Gegenüberstellung des CEM sowie verschiedener anderer Matchingverfahren findet sich bei Iacus/King/Porro (2011b).

⁶⁴ Vgl. Iacus/King/Porro (2011b), S. 356.

auf diese Weise 470 Beobachtungen (rund 99%) für die Analyse erhalten, in nur fünf Fällen fehlen entsprechende Beobachtungen in der Kontrollgruppe.⁶⁵ Durch das Matching erfolgt hier eine explizite Analyse des Effekts der Heirat auf die heiratenden Frauen, da all jene Beobachtungen der Kontrollgruppe unberücksichtigt bleiben, die kein Äquivalent in der Treatmentgruppe haben.

7 Ergebnisse

7.1 Coarsened Exact Matching

In den folgenden Abbildungen erfolgt zur Evaluation des Schätzmodells zunächst ein Ausweis der Ergebnisse von drei verschiedenen Modellspezifikationen: bei Berücksichtigung des vollen Satzes der Kontrollvariablen, bei Streichung der Persönlichkeitsmerkmale („ohne Big-5“) sowie ohne Kontrollvariablen. Es werden jeweils Schätzergebnisse unter Berücksichtigung robuster Standardfehler ausgewiesen.

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	-2.345* [-4.823, 0.133]	-2.329* [-4.855, 0.198]	-3.253** [-6.451, -0.0557]
Observations	3830	3830	3830
R-squared	0.429	0.406	0.010
Adjusted R-squared	0.423	0.402	0.010

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 6: CEM; sämtliche Beobachtungen.⁶⁶

Der signifikante Gesamteffekt⁶⁷ bei Berücksichtigung aller Kontrollen bzw. ohne die Big-5-Variablen beträgt -2,3 Stunden bzw. etwa -6,5% (Abbildung 6)⁶⁸. Da dieser Effekt durch die höhere Reduktion werdender Mütter getrieben sein mag, zeigen die folgenden Abbildungen die Ergebnisse einer differenzierteren Analyse, bei der bereits im Matching-Prozess nur diejenigen Frauen berücksichtigt wurden, die im Zeitraum von $t_{vor}+1$ bis t_{nach} nicht Mutter

⁶⁵ Auch von den Beobachtungen der Kontrollgruppe bleiben fast 96% (1.445 von 1.510) enthalten, da beim CEM alle einer Beobachtung der Treatmentgruppe entsprechenden Beobachtungen berücksichtigt werden. Von den Frauen aus der Kontrollgruppe (bzw. Treatmentgruppe) ohne Kind verbleiben 1.299 von 1.349 (bzw. sämtliche 311) sowie 146 von 161 (bzw. 159 von 164) der Frauen mit Kind.

⁶⁶ Darstellungen der Ergebnisse inkl. Ausweis sämtlicher Kontrollvariablen sind auf Anfrage verfügbar.

⁶⁷ Bei Betrachtung der Abbildungen sollte berücksichtigt werden, dass bei der angegebenen Beobachtungszahl erstens bereits die Gewichtungen der Kontrollgruppe berücksichtigt sind (bei der Treatmentgruppe betragen diese immer eins) und zweitens es hierdurch zu gerundeten Werten kommt.

⁶⁸ Die Werte im Fließtext beziehen sich, sofern nicht anders genannt, auf die Ergebnisse der Schätzungen ohne die Big-5-Variablen.

wurden (Abbildung 7) bzw. die in diesem Zeitraum ein Kind bekommen haben (Abbildung 8). Tatsächlich zeigt sich bei isolierter Betrachtung der Frauen mit Kind eine signifikante Reduktion um mehr als 6 Wochenstunden bzw. 15,9%. Im Fall, dass parallel zur Heirat kein Kind geboren wird, ist dagegen kein nennenswerter (und nicht signifikanter) Effekt durch die Heirat zu beobachten. Berücksichtigt man für repräsentative Ergebnisse die Gewichtungsfaktoren, die im SOEP zur Verfügung gestellt werden, reduzieren sich die geschätzten Arbeitsmarkteffekte in allen Gruppen. Der Gesamteffekt reduziert sich um mehr als die Hälfte auf etwa eine Wochenstunde (bzw. -2,8%). Dabei beträgt die Reduktion der Berufsstunden der ein Kind bekommenden Frauen nun nur noch 4,7 Stunden (-12,7%).⁶⁹ Grundsätzlich scheinen die Effekte bei Nichtberücksichtigung der Gewichtung überschätzt zu werden.

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	0.0150 [-2.558, 2.588]	0.0133 [-2.615, 2.642]	-0.0869 [-3.338, 3.164]
Observations	3220	3220	3220
R-squared	0.414	0.387	0.009
Adjusted R-squared	0.406	0.382	0.009

95% confidence intervals in brackets
 * p<0.1, ** p<0.05, *** p<0.01

Abbildung 7: CEM; nur Frauen, die im Zeitraum von $t_{vor} + 1$ bis t_{nach} kein Kind bekommen haben.

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	-6.132** [-11.34, -0.927]	-6.145** [-11.43, -0.865]	-9.447*** [-15.51, -3.379]
Observations	610	610	610
R-squared	0.523	0.499	0.223
Adjusted R-squared	0.488	0.476	0.220

95% confidence intervals in brackets
 * p<0.1, ** p<0.05, *** p<0.01

Abbildung 8: CEM; nur Frauen, die im Zeitraum von $t_{vor} + 1$ bis t_{nach} kein Kind bekommen haben.

Die Berücksichtigung der BIG-5-Variablen führt zu keiner bedeutsamen Beeinflussung der Ergebnisse – ein Indiz dafür, dass „unbeobachtbare“ Unterschiede zwischen beiden Gruppen trotz vereinzelter Signifikanz insgesamt in der Schätzung eine eher untergeordnete Rolle spielen –, ein Modell ohne Kontrollvariablen erscheint dagegen aufgrund der deutlich abweichenden Ergebnisse ungeeignet. Die folgenden Detailanalysen werden mit Blick auf die sinkende Beobachtungszahl nur auf Grundlage des „abgespeckten“ Modells ohne Big-5-Variablen

⁶⁹ Die Regressionsauswertungen zu den gewichteten Schätzungen finden sich zur besseren Übersicht im Anhang.

durchgeführt. Um auszuschließen, dass die – zumindest bei den Frauen, die kein Kind bekommen – nicht beobachtbare Reaktion des Arbeitsverhaltens auf die Heirat von „älteren“ Frauen, die bereits länger im Beruf stehen, verzerrt wird, wurde die gleiche Analyse auf Grundlage von Frauen, die maximal 35 Jahre alt sind, wiederholt. Die beobachteten Effekte sind hier aber sogar noch etwas geringer als vorher, in ihrer Größenordnung allerdings identisch.

Die bisherigen Ergebnisse widersprechen zwar nicht einer unter Umständen vorhandenen verstärkenden Wirkung der Heirat auf das Arbeitsangebot bei Frauen, die gleichzeitig ein Kind bekommen. Ohne die Geburt eines Kindes im gleichen Zeitraum ist eine Veränderung der Berufsstunden aber nicht belastbar festzustellen. Zumindest lässt sich daher sagen, dass allein die Heirat kurzfristig keine das Arbeitsverhalten der Frauen beeinflussende Wirkung entfaltet. Aufgrund der bereits weiter oben dargelegten Argumentation zu anderen mit der Heirat verbundenen Folgen, die potentiell das Arbeitsverhalten beeinflussen mögen, sprechen die Ergebnisse in der kurzen Frist daher auch gegen eine reale Auswirkung des Ehegattensplittings. Gleichzeitig zeigen die Ergebnisse aber auch – und dies ist eine zentrale Erkenntnis, auch bezogen auf das Forschungsdesign –, dass eine Anpassung der Berufsstunden in der kurzen Frist grundsätzlich beobachtbar ist. Diese Erkenntnis erscheint bezogen auf die Geburt eines Kindes trivial, nicht aber bezogen auf die messbar (deutlich) höhere Reduktion bei gleichzeitiger Heirat.

Da sich in den Ergebnissen zeigt, dass sich die Veränderung des Arbeitsverhaltens von Frauen abhängig von der Geburt eines Kindes fundamental unterscheidet und gemittelte Werte über den gesamten Datensatz daher weniger wertvolle Erkenntnisse liefern werden, erfolgen in der weiteren Analyse nur noch Ausweise zu den Untergruppen „ohne Geburt Kind“ sowie „nur Geburt Kind“. Da der steuerliche Effekt des Ehegattensplittings umso größer ist, 1.) je ungleicher die Einkommen verteilt sind sowie 2.) je höher das gemeinsame Einkommen ist, wird im nächsten Schritt nach diesen Merkmalen differenziert.⁷⁰ Theoretisch lässt ein größerer steuerlicher Effekt auch eine stärkere Reaktion bei den Berufsstunden erwarten. Frauen aus Haushalten mit einem höheren Gesamteinkommen sollten aus finanzieller Sicht außerdem eher in der Lage sein, ihr eigenes Arbeitseinkommen bzw. die gearbeiteten Stunden zu reduzieren. Für eine Differenzierung nach der Höhe des Haushaltseinkommens im Zeitpunkt t_{vor} wurden zunächst die

⁷⁰ Diese beiden Ebenen stehen auch im Zentrum der Diskussionen zum Ehegattensplitting; vgl. z. B. Buslei/Wrohlich (2014).

Beobachtungen beider Gruppen aus Haushalten mit Einkommen unterhalb sowie oberhalb des Medians⁷¹ getrennt – wie oben beschrieben – gematcht.⁷²

Obwohl große Effekte bei einer Wirkung des Ehegattensplittings aufgrund der höheren Grenzsteuersätze eher bei Frauen aus Haushalten mit hohen Einkommen zu erwarten sein sollten, ist das Gegenteil der Fall. Ein signifikanter Effekt ist nur bei Frauen aus Haushalten mit Einkommen unterhalb des Medians zu beobachten, falls diese ein Kind im Heiratszeitraum bekommen. Die der Heirat zuzurechnende Reduzierung von mehr als 11 Stunden entspricht dabei einer relativen Senkung der Berufsstunden um etwa ein Drittel (Abbildung 9). Bei Frauen aus Haushalten mit Einkommen oberhalb des Medians sind dagegen keine signifikanten Reaktionen zu messen. Selbst die der Heirat zuzurechnende Änderung bei Frauen mit Geburt eines Kindes ist nicht signifikant und liegt bei unter 3% im Vergleich zu den Stunden zum Zeitpunkt t_{vor} (Abbildung 10). So sprechen auch die Erkenntnisse dieser Differenzierung gegen eine generelle negative Wirkung des Ehegattensplittings.

	ohne Geburt Kind	nur Geburt Kind
DiD-Schätzer	-0.0242 [-4.474, 4.425]	-11.56*** [-18.80, -4.314]
Observations	1434	354
R-squared	0.360	0.473
Adjusted R-squared	0.348	0.431

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 9: CEM; nur Frauen aus Haushalten mit Einkommen unterhalb des Medians (zum Zeitpunkt t_{vor}).

⁷¹ Bezogen auf die Haushaltsnettoeinkommen der heiratenden Frauen vor dem Matching zum Zeitpunkt t_{vor} . Frauen aus Haushalten mit Einkommen auf Medianniveau wurden allesamt der unterhalb-des-Medians-Gruppe zugeordnet. Die Berechnung des Medians erfolgte bezogen auf die Frauen der Treatmentgruppe und vor Differenzierung zwischen Frauen mit und ohne Kindern.

⁷² In der „Unter-Median-Gruppe“ bleiben je Zeitpunkt 234 der 241 Beobachtungen aus der Treatment- und 660 der 720 Beobachtungen aus der Kontrollgruppe nach dem Matching erhalten. In der „Über-Median-Gruppe“ sind es jeweils 233 der 234 Beobachtungen der Treatment- und 724 der 790 Beobachtungen aus der Kontrollgruppe. Insbesondere die Beobachtungen der heiratenden Frauen bleiben also fast vollständig erhalten, so dass annähernd der allgemeine ATET gemessen werden kann.

	ohne Geburt Kind	nur Geburt Kind
DiD-Schätzer	-0.490 [-3.576, 2.595]	-1.183 [-9.088, 6.722]
Observations	1672	242
R-squared	0.395	0.605
Adjusted R-squared	0.385	0.557

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 10: CEM; nur Frauen aus Haushalten mit Einkommen oberhalb des Medians (zum Zeitpunkt t_{vor}).

Betrachtet man die Veränderung der Berufsstunden abhängig vom Einkommensanteil der Frau bezogen auf das gemeinsame Einkommen innerhalb der Partnerschaft⁷³ sollte man, glaubt man den Befürchtungen der Kritiker des Ehegattensplittings, vor allem Reaktionen bei Frauen als Zweitverdienern finden.⁷⁴ Als Einkommensanteil der Frau gilt für diese Abgrenzung der Anteil des selbst erwirtschafteten Bruttoeinkommens am gesamten Bruttoeinkommen des Haushalts.⁷⁵ Dabei erfolgt eine Unterteilung in drei Gruppen: Frauen, die 1.) maximal ein Drittel des gemeinsamen Einkommens (bei denen sich das Ehegattensplitting also auswirkt) (Abbildung 11), 2.) über ein Drittel, aber weniger als zwei Drittel des gemeinsamen Einkommens (Abbildung 12) sowie 3.) mindestens zwei Drittel des gemeinsamen Einkommens erwirtschaften (kein Splittingeffekt) (Abbildung 13).⁷⁶ Eine weitere Unterteilung anhand des Einkommensanteils der Frau – die insbesondere für die Frauen der erstgenannten Gruppe, die Kinder bekommen, interessant wäre – ist aufgrund des begrenzten Beobachtungsumfanges nicht sinnvoll. In der

⁷³ Auch hier bezieht sich der Einkommensanteil jeweils auf den Zeitpunkt t_{vor} .

⁷⁴ Vgl. z. B. Buslei/Wrohlich (2014), S. 4, sowie die dort wie auch in dieser Arbeit genannten empirischen Arbeiten zu Arbeitsmarkteffekten des Ehegattensplittings wie z. B. von Steiner/Wrohlich (2004), Bach et al. (2011), Bonin et al. (2013) oder Müller et al. (2013).

⁷⁵ Zweckmäßiger hinsichtlich einer Abgrenzung der auf die Berufsstunden der Partner bezogenen Einkommen wäre u. U. das Verhältnis des genannten Einkommens der Frau zum entsprechenden Einkommen des Mannes. Allerdings fehlen hierzu bei einem größeren Teil der männlichen Partner Angaben, so dass bei dieser Betrachtung nur 389 der 475 beobachteten heiratenden Frauen berücksichtigt werden könnten. Die alternative Berechnung des Anteils am Haushaltseinkommen scheint zudem aufgrund der ähnlichen Ausprägungen (arithmetisches Mittel bei 40,73% (Arbeitseinkommen Frau/Haushaltseinkommen) zu 40,07% (Arbeitseinkommen Frau/Arbeitseinkommen beide Partner) bzw. einem Korrelationskoeffizienten beider Werte von über 0,86 ein geeignetes Maß zu sein.

⁷⁶ Nach dem Matching bleiben je Zeitpunkt 153 von 177 bzw. 490 von 530 Beobachtungen der Treatment- bzw. Kontrollgruppe (Einkommensanteil der Frau maximal 1/3) sowie 230 von 232 bzw. 685 von 756 Beobachtungen (Einkommensanteil der Frau über 1/3, aber unter 2/3) enthalten. Von den Frauen mit einem Einkommensanteil von mindestens 2/3 bleiben – hier werden nur Frauen ohne Nachwuchs betrachtet – nach dem Matching 42 von 42 bzw. 146 von 202 Beobachtungen der Treatment- bzw. Kontrollgruppe erhalten.

letzten genannten Gruppe liegt zudem die Zahl der gebährenden Frauen in der Treatmentgruppe im einstelligen Bereich, weshalb hierzu keine Schätzung erfolgt.

	ohne Geburt Kind	nur Geburt Kind
DiD-Schätzer	3.682 [-0.801, 8.166]	-9.951 [-22.18, 2.277]
Observations	1136	150
R-squared	0.439	0.455
Adjusted R-squared	0.426	0.340

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 11: CEM; nur Frauen, die im Zeitpunkt t_{vor} max. 1/3 des gemeinsamen Einkommens erwirtschaftet haben.

	ohne Geburt Kind	nur Geburt Kind
DiD-Schätzer	-2.773* [-5.847, 0.302]	-5.426 [-12.35, 1.500]
Observations	1496	334
R-squared	0.234	0.559
Adjusted R-squared	0.220	0.522

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 12: CEM; nur Frauen, die im Zeitpunkt t_{vor} mehr als 1/3, aber weniger als 2/3 des gemeinsamen Einkommens erwirtschaftet haben.

	ohne Geburt Kind
DiD-Schätzer	-1.935 [-8.689, 4.819]
Observations	376
R-squared	0.269
Adjusted R-squared	0.214

95% confidence intervals in brackets

* p<0.1, ** p<0.05, *** p<0.01

Abbildung 13: CEM; nur Frauen, die im Zeitpunkt t_{vor} mind. 2/3 des gemeinsamen Einkommens erwirtschaftet haben.

Die in den Abbildungen 11 bis 13 aufgeführten Ergebnisse widersprechen den genannten Aussagen der Splittingkritiker, dass insbesondere Frauen in der Zweitverdiener-Rolle aufgrund des Ehegattensplittings ihre Berufstätigkeit reduzieren würden. Es scheint sogar das Gegenteil

der Fall zu sein: Betrachtet man nur Frauen, die im Beobachtungszeitraum kein Kind bekommen, lässt sich gerade bei den Frauen, die maximal 1/3 des Haushaltseinkommen erwirtschaften, sogar eine leicht positive Entwicklung der Berufsstunden im Zuge der Heirat feststellen. Sehr deutlich ist dagegen der Befund, sofern gleichzeitig ein Kind geboren wird. Hier kommt es bei den heiratenden Frauen, die maximal 1/3 des gesamten Einkommens erwirtschaften, zu einer um fast zehn Stunden höheren Senkung der wöchentlich für den Beruf aufgewendeten Zeit. Dies entspricht einem relativen Effekt von fast 40%. Zwar ist die Zahl der bei dieser Schätzung berücksichtigten Frauen relativ klein – ein möglicher Grund für die fehlende Signifikanz. Eine Wirkung des Ehegattensplittings kann aufgrund des deutlichen Heiratseffekts für diese Gruppe allerdings auf keinen Fall ausgeschlossen werden (Abbildung 11). Bei Frauen mit einem Anteil von mindestens 2/3 am gemeinsamen Einkommen kommt es zu einer heiratsinduzierten, jedoch nicht signifikanten Senkung der Berufsstunden von 1,9 Stunden bzw. 4,7% (Abbildung 13). In Beziehungen mit ähnlichen Einkommensanteilen beider Partner kommt es sogar zu noch höheren Absenkungen der Berufsstunden der Frau aufgrund der Heirat. Diese Senkung ist deutlicher bei Frauen mit Nachwuchs (-5,4 Stunden bzw. -12,5%), jedoch gegenüber der von Frauen ohne Kind (-2,8 Stunden bzw. -6,2%) nicht signifikant⁷⁷ (Abbildung 12). Beachtlich an den nach Einkommensanteilen differenzierten Ergebnissen ist, dass sich die bisher im Mittel nicht messbare Wirkung der Heirat auf das Arbeitsverhalten von Frauen ohne Kinder aus einem positiven Effekt auf Seiten der „Zweitverdiener-Frauen“ (Einkommensanteil bis 1/3) und einem negativen Effekt der restlichen Frauen zusammensetzt. Dieser Befund widerspricht zwar einer Wirkung des Ehegattensplittings. Jedoch sollte auf weitere Spekulationen verzichtet werden, auch da die Ergebnisse z. T. aufgrund geringer Beobachtungszahlen nicht vollends belastbar scheinen.

7.2 Validierung der Ergebnisse

Zunächst sollen die im vorherigen Abschnitt geschätzten Effekte durch eine Regression über alle Daten ohne Matching validiert werden. Wie in den vorhergehenden Analysen bleiben hierfür die Big-5-Variablen unberücksichtigt, alle anderen Charakteristika gehen allerdings als unabhängige Variablen in die Regression ein. Bei Betrachtung des gesamten ungewichteten Datensatzes ergeben sich ähnliche Ergebnisse: Die Reduzierung der Berufsstunden bei den Frauen, die im gleichen Zeitraum ein Kind bekommen, beträgt hier 6,0 Stunden und bleibt wie der relative Wert (-15,5%) im Vergleich zu den Ergebnissen des CEM annähernd identisch. Bei Frauen ohne Nachwuchs kommt es in Folge der Heirat wieder zu einem minimalen, wenngleich nicht signifikanten Anstieg der Berufsstunden.

⁷⁷ Bei Berücksichtigung der Gewichtung verschwindet die genannte Signifikanz allerdings.

Für eine weitere Validierung der Ergebnisse soll der Effekt der Heirat auf die Berufsstunden auf Grundlage eines Propensity Score Matchings geschätzt werden. Hierzu werden Frauen der Treatment- und Kontrollgruppe auf Basis ihrer Heiratswahrscheinlichkeiten zum Zeitpunkt t_{vor} gegeben persönlicher Charakteristika gematcht und anschließend jeweils die Veränderungen der Berufsstunden verglichen. Berücksichtigt werden diejenigen der oben aufgeführten Charakteristika zum Zeitpunkt t_{vor} , bei denen ein Zusammenhang zur Heirat festgestellt werden kann. Dies sind das Alter, die Religiosität, die Bildung, der Ost-Dummy, der Stadt-Dummy, die Anzahl der Kinder, das Vorhandensein eines jungen Kindes (unter 6 Jahre), die Jahre im Beruf, der Stundenlohn, das sonstige Haushaltseinkommen, das eigene Vermögen bzw. das des Partners sowie die berufsbezogenen Dummies zur Betriebsgröße sowie dem öffentlichen Dienst. Außerdem wird zusätzlich noch die Dauer der Beziehung in Jahren berücksichtigt.

Betrachtet man diejenigen Frauen ohne die Geburt eines Kindes im Zeitraum der potentiellen Heirat ergibt sich ein nicht signifikanter, aber im Vergleich zu den vorherigen Ergebnissen etwas höherer ATET von -1,1 Wochenstunden. Der Heiratseffekt bei gleichzeitiger Geburt eines Kindes liegt dagegen wie vorher bei etwa 5,5 Wochenstunden, ist hier jedoch ebenfalls nicht signifikant. Beide Ergebnisse liegen in einer ähnlichen Größenordnung wie die der vorherigen Analysen und können die bisherigen Schätzungen daher bestätigen.

	ohne Geburt Kind	nur Geburt Kind
ATET	-1.069 [-4.313, 2.175]	-5.518 [-14.29, 3.256]
Observations	1660	325

95% confidence intervals in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Abbildung 14: Ergebnisse des Propensity Score Matchings.

8 Schlussfolgerungen

Die vorliegende Arbeit hat zum Ziel, einen Beitrag zu der für die Bewertung der Ehegattenbesteuerung elementaren Frage zu leisten, ob das Ehegattensplitting Frauen dazu veranlasst, weniger zu arbeiten. Obwohl die bisherige Diskussion in der Regel daran kaum Zweifel aufkommen lässt, ergibt sich durch die in dieser Arbeit zusammengetragenen Erkenntnisse ein differenzierteres Bild. Ohne die Geburt eines Kindes in den Jahren um die Heirat kommt es im Vergleich zu in NEL lebenden Frauen im gleichen Zeitraum zu keinen bzw. nicht signifikant

abweichenden Reduzierungen der Berufsstunden bei heiratenden Frauen. Lediglich bei gleichzeitiger Geburt von Kindern lässt sich auch in der vorliegenden Studie beobachten, dass Frauen ihr Arbeitsverhalten nicht nur aufgrund des Nachwuchses anpassen, sondern dies im Falle einer Eheschließung (deutlich) verstärkt tun. Kurzfristige Reaktionen sind also vorhanden und auch messbar. Die betroffenen Frauen scheinen ihr Arbeitsverhalten allerdings nicht allgemein aufgrund der Heirat und daher insbesondere nicht aufgrund einer „Anreizwirkung“ des Ehegattensplittings zu tun. Die beobachtbare Reaktion im Zusammenhang mit der Geburt von Kindern könnte möglicherweise auf die Wirkung des Versicherungseffekts der Heirat sein, wenngleich die vorliegende Arbeit den sicheren Bezug zu einer Ursache nicht zulässt.

Es sollte allerdings bei der Bewertung dieses Befunds und insbesondere im Vergleich zu Ergebnissen anderer Arbeiten bedacht werden, dass es in den letzten Jahrzehnten einen Anstieg der Arbeitsmarktpartizipation von Frauen und einen einhergehenden Bedeutungsverlust traditioneller Rollenbilder gegeben hat. Gleichzeitig wurden in dieser Arbeit *heiratende* (also nicht allgemein verheiratete) Frauen mit einem bezogen auf die Gesamtbevölkerung niedrigen Durchschnittsalter von etwas über 30 Jahren betrachtet. Es lassen sich daher insbesondere keine Rückschlüsse ziehen, inwiefern den verheirateten Frauen in Deutschland in ihrer Gesamtheit durch das Ehegattensplitting oder allgemein die Heirat Anreize, zu arbeiten, genommen wurden und werden. Bezogen auf die „aktuell heiratende Generation“ scheint es allerdings keine generellen Effekte der Heirat zu geben. Es kann daher geschlussfolgert werden, dass es für heutzutage (auch in den Arbeitsmarkt eintretende) Frauen unter Umständen andere (bedeutendere) Gründe gibt, sich für oder gegen eine Karriere oder allgemein eine Erwerbsbeteiligung zu entscheiden als den (sich bei tendenziell angleichenden Einkommen innerhalb einer Partnerschaft sinkendem) „Vorteil“ des Ehegattensplittings. Für die Entwicklung sich angleichender Einkommen sorgt neben der steigenden Partizipation von Frauen am Arbeitsmarkt beispielsweise auch das bereits ausgiebig diskutierte⁷⁸ und aktuell für Deutschland unter anderem von Triebe (2015a) untersuchte selektive Verhalten bei der Partnerwahl („Assortative Mating“): Männer und Frauen suchen danach immer mehr nach sich selbst ähnelnden Partnern, insbesondere auch bezogen auf sozioökonomische Merkmale wie der Bildung und dem Einkommen.

Ein zweiter elementarer Befund dieser Arbeit ist, dass gerade Frauen, die einen geringen Anteil zum Haushaltseinkommen beitragen sowie allgemein Frauen aus einkommensschwächeren Haushalten ihre Erwerbstätigkeit nach der Eheschließung im Vergleich zu in NEL lebenden Frauen deutlich stärker reduzieren, wenn sie gleichzeitig ein Kind bekommen. Tatsächlich scheint hier ein Zusammenhang zwischen der Heirat und der Reduzierung der Arbeitsstunden zu

⁷⁸ Siehe hierzu bspw. Becker (1981).

bestehen. Ob das Ehegattensplitting eine zentrale Rolle spielt, lässt sich an dieser Stelle weder ausschließen noch bestätigen. Es ist denkbar, dass gerade die Frauen, die ihre Berufstätigkeit ohnehin hinter die Betreuung ihrer Kinder stellen wollen und daher freiwillig reduzieren, die „Vorteile“ der Heirat (inkl. des Ehegattensplittings) nutzen und daher vermehrt in der Treatmentgruppe vertreten sind.⁷⁹ Die Richtung der Wirkung mag aber auch umgekehrt sein. werdende Mütter könnten ihre Berufsstunden gerade aufgrund der erheblichen finanziellen Absicherung durch eine nicht von diesen Motiven getriebene Heirat reduzieren.

⁷⁹ Hierfür sprechen z. B. die Ergebnisse von Barg/Beblo (2012).

Adamopoulou, Effrosyni (2010) Will You „Quasi-marry“ Me? The Rise of Cohabitation and Decline of Marriages; Working Paper, Economic Series 10-26, Universidad Carlos III de Madrid.

Allport, Gordon W. und Henry S. Odbert (1936) Trait-Names – A Psycho-lexical Study; in: Psychological Review Publications, Vol. 47, Nr. 1.

Angrist, Joshua D. und Alan B. Krueger (1999) Empirical Strategies in Labor Economics; in: Ashenfelter, Orley C. und David Card (Hrsg.) Handbook of Labor Economics; Vol. 3, Part A, S. 1277 – 1366.

Asef, Dominik, Susanne Wanger und Ines Zapf (2011) Statistische Messung des Arbeitseinsatzes; in: WiSta 2011, S. 1058 – 1064.

Asendorf, Jens B. und Franz J. Neyer (2012) Psychologie der Persönlichkeit; 5. Auflage, Springer.

Bach, Stefan, Johannes Geyer, Peter Haan und Katharina Wrohlich (2011) Reform des Ehegattensplittings: Nur eine reine Individualbesteuerung erhöht die Erwerbsanreize deutlich; DIW Wochenbericht 41/2011, S. 13 – 19.

Barg, Katherin und Miriam Beblo (2012) Does „Sorting into Specialization“ Explain the Differences in Time Use between Married and Cohabiting Couples? An Empirical Application for Germany; in: Annals of Economics and Statistics, Nr. 105/106, S. 127 – 152.

Bargain, Olivier, Kristian Orsini und Andreas Peichl (2014) Comparing Labor Supply Elasticities in Europe and the United States – New Results; in: The Journal of Human Resources; Vol. 49, Nr. 3, S. 723 – 838.

Baumgarten, Jörg und Henriette Houben (2014) Die Erwerbs- und Verbrauchsgemeinschaft von Ehepaaren: Eine empirische Studie; in: Steuer und Wirtschaft, 2/2014, S. 116 – 131.

Becker, Gary S. (1981) A Treatise on the Family; Harvard University Press.

Beninger, Denis, Francois Laisney und Miriam Beblo (2007) Welfare Analysis of a Tax Reform for Germany: A Comparison of the Unitary and Collective Models of Household Labour Supply; in: Journal of Population Economics, Vol. 20, Nr. 4, S. 869 – 893.

Blau, Francine D. und Lawrence M. Kahn (2007) Changes in the Labor Supply Behavior of Married Women: 1980 – 2000; in: Journal of Labor Economics, Vol. 25, Nr. 3, S. 393 - 438.

Blaufus, Kay, Jonathan Bob, Jochen Hundsdoerfer, Christian Sielaff, Dirk Kiesewetter und Joachim Weimann (2015) Perception of income tax rates: evidence from Germany; in: European Journal of Law and Economics, Vol. 40, Nr. 3, S. 457 – 478.

Bonin, Holger, Markus Clauss, Irene Gerlach, Inga Laß, Anna L. Mancini, Marc-André Nehr Korn-Ludwig, Verena Niepel, Reinhold Schnabel, Holger Stichnoth und Katharina Sutter (2013) Evaluation zentraler ehe- und familienbezogener Leistungen in Deutschland; ZEW, finales Gutachten im Auftrag der Prognos AG.

Borkenau, Peter und Fritz Ostendorf (2008) NEO-FFI – NEO-Fünf-Faktoren-Inventar nach Costa und McCrae; 2. Auflage, Hogrefe.

Buslei, Hermann und Katharina Wrohlich (2014) Besteuerung von Paaren: das Ehegattensplitting und seine Alternativen; DIW Roundup 21.

Cameron, A. Colin und Pravin K. Trivedi (2005) Microeconometrics – Methods and Applications; Cambridge University Press.

- Cornelißen, Waltraud (2005) Gender Datenreport – 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland; München, November 2005.
- Crossley, Thomas F. und Sund-Hee Jeon (2007) Joint Taxation and the Labour Supply of Married Women: Evidence from the Canadian Tax Reform of 1988; in: Fiscal Studies, Vol. 28, Nr. 3, S. 343 – 365.
- Dehne, Max und Jürgen Schupp (2007) Persönlichkeitsmerkmale im Sozio-oekonomischen Panel (SOEP) – Konzept, Umsetzung und empirische Eigenschaften; DIW Berlin Research Notes, Nr. 26.
- Eissa, Nada (2002) Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment; Working Paper, University of California, Berkeley and NBER.
- Enders, Craig K. (2010) Applied Missing Data Analysis; The Guilford Press, New York.
- Fehr, Hans, Manuel Kallweit und Fabian Kindermann (2015) Reforming Family Taxation in Germany – Labor Supply versus Insurance Effects; in: FinanzArchiv: Public Finance Analysis, Vol. 71, Nr. 1, S. 53 – 81.
- Fitzenberger, Bernd, Katrin Sommerfeld und Susanne Steffes (2013) Causal effects on employment after first birth – A dynamic treatment approach; in: Labour Economics, Vol. 25, S. 49 – 62..
- Frederiksson, Peter und Per Johansson (2008) Dynamic Treatment Assignment: The Consequences for Evaluations Using Observational Data; in: Journal of Business & Economic Statistics; Vol. 26, Nr. 4, S. 435 – 445.
- Gerlitz, Jean-Yves und Jürgen Schupp (2005) Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale des SOEP; DIW Berlin Research Notes 4, Juli 2005.
- Grziwotz, Herbert (2014) Nichteheleiche Lebensgemeinschaft; NJW Praxis Band 39, 5. Auflage, C.H. Beck, München.
- Gustafsson, Siv (1992) Separate taxation and married women's labor supply. A comparison of West Germany and Sweden; in: Journal of Population Economics, Vol. 5, Nr. 1, S. 61 – 85.
- Iacus, Stefano M., Gary King und Giuseppe Porro (2011a) Causal Inference without Balance Checking: Coarsened Exact Matching; in: Political Analysis, Oxford University Press on behalf of the Society for Political Methodology.
- Iacus, Stefano M., Gary King und Giuseppe Porro (2011b) Multivariate Matching Methods That Are Monotonic Imbalance Bounding; in: Journal of the American Statistical Association, Vol. 106, Nr. 493, S. 345 – 361.
- Iacus, Stefano M., Gary King und Giuseppe Porro (2015) A Theory of Statistical Inference for Matching Methods in Applied Causal Research; Working Paper.
- Keane, Michael P. (2010) Structural vs. atheoretic approaches to econometrics; in: Journal of Econometrics, 156, S. 3 – 20.
- Kempe, Wolfram (1996) Das Arbeitsangebot verheirateter Frauen in den neuen und alten Bundesländern – Eine semiparametrische Regressionsanalyse; Statistische Diskussionsbeiträge, Nr. 2, Universität Potsdam..
- Krzepkowski, Matt und Jack Mintz (2013) No more second-class taxpayers: How income splitting can bring fairness to Canada's single-income families; SPP Research Paper, University of Calgary, Vol. 6, Issue 14.
- Kumar, Anil und Che-Yuan Liang (2016) Declining Female Labor Supply Elasticities in the United States and Implications for Tax Policy: Evidence from Panel Data; in: National Tax Journal, Vol. 69, Nr. 3, S. 481 – 516.

- Laczó, Sarolta (2011) Living Arrangements and Labour Supply: Evidence from the United Kingdom; Working Paper der University of California, 31.
- LaLumia (2008) The Effects of Joint Taxation of Married Couples on Labor Supply and Non-wage Income; in: Journal of Public Economics, Vol. 92, Nr. 7, S. 1698 – 1719.
- Maiterth, Ralf und Malte Chirvi (2015) Das Ehegattensplitting aus Sicht der Steuerwissenschaften; in: Steuer und Wirtschaft, Nr. 1, S. 19 – 32.
- Mincer, Jacob A. (1974) Schooling, Experience and Earnings; NBER Books.
- Müller, Kai-Uwe, C. Katharina Spieß, Chrysanthi Tsiasioti, Katharina Wrohlich, Elisabeth Bügelmayer, Luke Haywood, Frauke Peter, Marko Ringmann und Sven Witzke (2013) Evaluationsmodul: Forderung und Wohlergehen von Kindern; Politikberatung kompakt 73.
- OECD (2015) OECD Family Database; <http://www.oecd.org/els/family/database.htm>.
- Puhani, Patrick A. (1995) Labour Supply of Married Women in Poland: A Microeconomic Study Based on the Polish Labour Force Survey; Discussion Paper No. 95-12, Zentrum für Europäische Wirtschaftsforschung.
- Rust, John (2010) Comments on: „Structural vs. atheoretic approaches to econometrics“ by Michael Keane; in: Journal of Econometrics, Vol. 156, S. 21 – 24.
- Sausgruber, Rupert und Hannes Winner (1996) Die Familie in den „drei Steuerwissenschaften“ – Der Versuch einer Systematisierung gegensätzlicher Positionen; in: Thöni, Erich und Hannes Winner (Hrsg.) Die Familie im Sozialstaat – Familienbesteuerung aus ökonomischer und juristischer Sicht; Veröffentlichungen der Universität Innsbruck Nr. 216, S. 241 – 274.
- Schröder, Melanie und Norma Burow (2016) Couple`s Labor Supply, Taxes, and the Division of Housework in a Gender-Neutral Lab; SOEP Discussion Paper 1593.
- Schupp, Jürgen und Jean-Yves Gerlitz (2014) Big Five Inventory-SOEP (BFI-S); Zusammenstellung sozialwissenschaftlicher Items und Skalen des gesis (Leibniz-Institut für Sozialwissenschaften).
- Selin, Håkan (2014) The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971; in: International Tax and Public Finance; Vol. 21, S. 894 – 922.
- Sianesi, Barbara (2004) An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s; in: The Review of Economics and Statistics; Vo. 86, Nr. 1, S. 133 – 155.
- Spahn, Paul B., Helmut Kaiser und Thomas Kassella (1992) The Tax Dilemma of Married Women in Germany; in: Fiscal Studies, Vol. 13, Nr. 2, S. 22 – 47.
- Steiner, Viktor und Katharina Wrohlich (2004) Household Taxation, Income Splitting and Labor Supply Incentives: A Microsimulation Study for Germany; SOEPPaper Nummer 421.
- Steiner, Viktor und Katharina Wrohlich (2008) Introducing Family Tax Splitting in Germany: How Would It Affect The Income Distribution, Work Incentives and Household Welfare?; in: FinanzArchiv: Public Finance Analysis, Vol 64, Nr. 1, S. 115 – 142.
- Triebe, Doreen (2015) To Marry or Not to Marry – Essays on Partnership Formation and Economic Labor Market Behavior of Married and Cohabiting Couples; Dissertation an der Technischen Universität Berlin.
- Triebe, Doreen (2015a) Like or Dislike? – Assortative Mating of Married and Cohabiting Couples & its Consequences on Income Inequality; in: Triebe, Doreen (2015) To Marry or Not to Marry – Essays on Partnership Formation and Economic Labor Market Behavior of Married and Cohabiting Couples; Dissertation an der Technischen Universität Berlin, S. 8 – 36.

Triebe, Doreen (2015b) The Added Worker Effect Differentiated by Gender and Partnership – Evidence from Involuntary Job Loss Status; in: Triebe, Doreen (2015) To Marry or Not to Marry – Essays on Partnership Formation and Economic Labor Market Behavior of Married and Cohabiting Couples; Dissertation an der Technischen Universität Berlin, S. 37 - 63.

Triebe, Doreen (2015c) Women at Work? The Impact of Cohabiting and Married Partners' Earnings on Work Hours; in: Triebe, Doreen (2015) To Marry or Not to Marry – Essays on Partnership Formation and Economic Labor Market Behavior of Married and Cohabiting Couples; Dissertation an der Technischen Universität Berlin, S. 64 - 100.

Van Soest, Arthur (1995) Structural Models of Family Labor Supply: A Discrete Choice Approach; in: Journal of Human Resources, Vol. 30, Nr. 1, S. 63 – 88.

Wagenhals, Gerhard (2008) Auswirkungen einer Reform des Ehegattensplittings; in: Seel, Barbara (Hrsg.) Ehegattensplitting und Familienpolitik; VS Verlag für Sozialwissenschaften, S. 239 – 267.

Wagner, Gert G., Joachim R. Frick und Jürgen Schupp (2007) The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements; SOEPpapers, Nr. 1.

Wagner, Gert G., Jan Göbel, Peter Krause, Rainer Pischner und Ingo Sieber (2008) Das Sozio-oekonomische Panel (SOEP): Multidisziplinäres Haushaltspanel und Kohortenstudie für Deutschland; in: AStA Wirtschafts- und Sozialstatistisches Archiv, Vol. 2, Nr. 4, S. 301 – 328.

Wolf, Elke (2010) Lohndifferenziale zwischen Vollzeit- und Teilzeitbeschäftigten in Ost- und Westdeutschland; WSI-Diskussionspapier Nr. 174.

Wolpin, Kenneth I. (2013) The Limits of Inference without Theory; MIT University Press, 2013.

Wrohlich, Katharina A. (2007) Evaluating Family Policy Reforms Using Behavioral Microsimulation; Dissertation an der Freien Universität Berlin.

10 Anhang

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	-1.022 [-4.864, 2.821]	-0.943 [-4.762, 2.877]	-1.616 [-9.474, 6.242]
Observations	3830	3830	3830
R-squared	0.497	0.479	0.028

95% confidence intervals in brackets
 * p<0.1, ** p<0.05, *** p<0.01

Anhang 1: Gewichtete Ergebnisse zu Abbildung 6: CEM; alle Frauen.

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	1.627 [-1.728, 4.982]	1.606 [-1.757, 4.968]	0.703 [-3.577, 4.983]
Observations	3220	3220	3220
R-squared	0.391	0.372	0.022

95% confidence intervals in brackets
 * p<0.1, ** p<0.05, *** p<0.01

Anhang 2: Gewichtete Ergebnisse zu Abbildung 7: CEM; nur Frauen, die im Zeitraum von $t_{vor} + 1$ bis t_{nach} kein Kind bekommen haben.

	volle Kontrollen	ohne Big-5	ohne Kontrollen
DiD-Schätzer	-4.890 [-12.10, 2.323]	-4.662 [-12.19, 2.862]	-4.561 [-16.46, 7.337]
Observations	610	610	610
R-squared	0.650	0.617	0.247

95% confidence intervals in brackets
 * p<0.1, ** p<0.05, *** p<0.01

Anhang 3: Gewichtete Ergebnisse zu Abbildung 8: CEM; nur Frauen, die im Zeitraum von $t_{vor} + 1$ bis t_{nach} ein Kind bekommen haben.

Impressum:

Arbeitskreis Quantitative Steuerlehre, arqus, e.V.

Vorstand: Prof. Dr. Ralf Maiterth (Vorsitzender),
Prof. Dr. Kay Blaufus, Prof. Dr. Dr. Andreas Löffler
Sitz des Vereins: Berlin

Herausgeber: Kay Blaufus, Jochen Hundsdoerfer,
Martin Jacob, Dirk Kiesewetter, Rolf J. König,
Lutz Kruschwitz, Andreas Löffler, Ralf Maiterth,
Heiko Müller, Jens Müller, Rainer Niemann,
Deborah Schanz, Sebastian Schanz, Caren Sureth-
Sloane, Corinna Treisch

Kontaktadresse:

Prof. Dr. Caren Sureth-Sloane, Universität Paderborn,
Fakultät für Wirtschaftswissenschaften,
Warburger Str. 100, 33098 Paderborn,
www.arqus.info, Email: info@arqus.info

ISSN 1861-8944