

Beeinflussen medizinische und vergütungsrelevante Schwellenwerte die Krankenhausbehandlung?

Eine empirische Analyse der Neugeborenenversorgung mit der
DRG-Statistik

Sebastian Wichert ¹²³

¹ifo Institut - Leibniz-Institut für Wirtschaftsforschung an der Universität
München e.V.

²Ludwig-Maximilians-Universität München

³Münchener Zentrum für Gesundheitswissenschaften (MC-Health)

Statistiktage 2018

20. Juli 2018

Motivation I

- ▶ Geringes Geburtsgewicht (GG) → schlechtere Gesundheit (z.B. Hack et al. 2002, Hummer et al. 2014).
- ▶ GG unterhalb best. Schwellenwert (aus med. Leitlinien) kann zu veränderter Behandlungsintensität und -qualität führen (z.B. Almond et al. 2010).
- ▶ Kleine GG-Änderungen können große Vergütungsänderungen mit sich bringen. → Geringes GG wird berichtet ("Upcoding") (z.B. Jürges and Köberlein 2015; Shigeoka and Fushimi 2014).
- ▶ Ein berichtetes geringes GG könnte daher positiv wirken:
 - ▶ Höhere Vergütung → mehr (teurere) Versorgung
 - ▶ Medizinische Leitlinien → intensivere Versorgung

Motivation II

Ausgangspunkt: Ob ein berichtetes GG unterhalb einer spezifischen Schwelle im dt. Fallpauschalensystem (G-DRG) zu unterschiedlicher Behandlung (Intensität/Qualität) führt, ist eine empirische Frage, der wir in unserem Projekt nachgegangen sind.

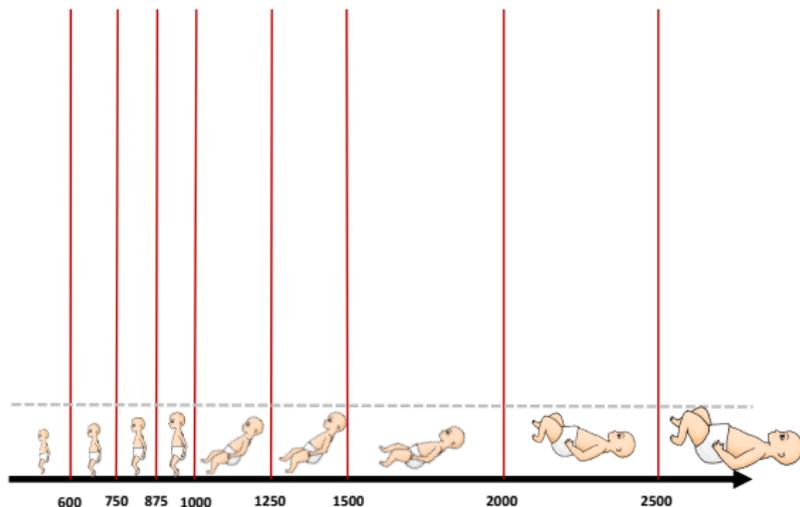
Grundlagenpapier: Reif, Simon; Wichert, Sebastian; und Amelie Wuppermann (2018): Is it good to be too light? Consequences of birth weight thresholds in hospital reimbursement systems, *Journal of Health Economics*, 59, 01-25.

Vorgehen:

- ▶ Dokumentation von Unterschieden bei der Behandlung und der Mortalität um GG-Schwellen im dt. G-DRG-System unter Verwendung administrativer Abrechnungsdaten aller Krankenhausgeburten in Deutschland von 2005-2011.
- ▶ Berücksichtigung der Risikoselektion bei der ökonomischen Schätzung führt zum Verschwinden der Behandlungsunterschiede.

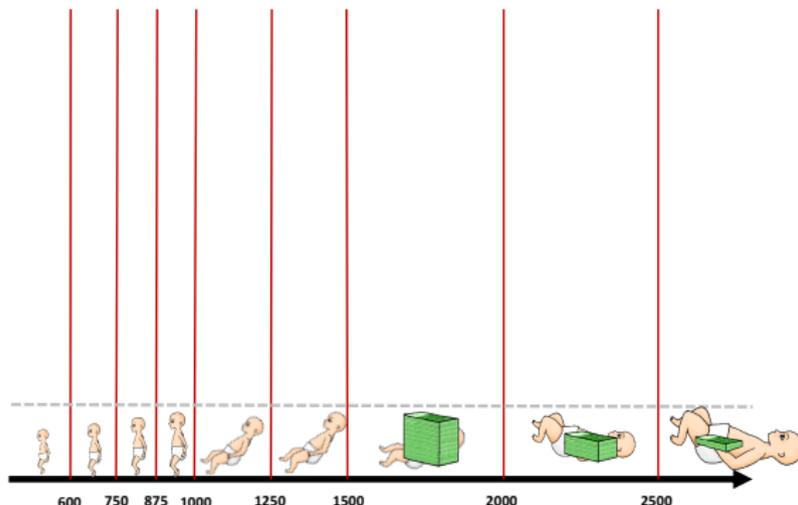
Deutsches Fallpauschalensystem (G-DRG) I

- ▶ Über 1000 Fallgruppen (DRGs), davon 42 DRGs für Neugeborenenbehandlung, davon 37 mit GG.
- ▶ 8 Geburtsgewichtsschwellen (X-Achse) + bis zu 6 Schweregrade (Y-Achse)



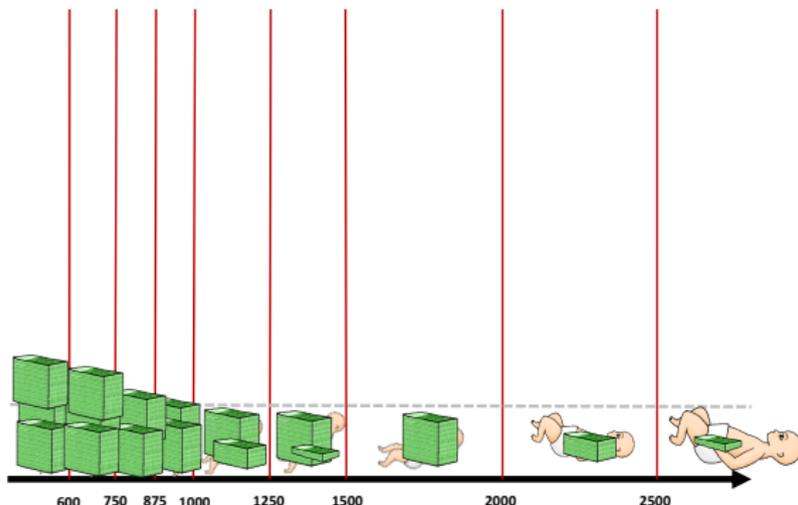
Deutsches Fallpauschalensystem (G-DRG) I

- ▶ Über 1000 Fallgruppen (DRGs), davon 42 DRGs für Neugeborenenbehandlung, davon 37 mit GG.
- ▶ 8 Geburtsgewichtsschwellen (X-Achse) + bis zu 6 Schweregrade (Y-Achse)



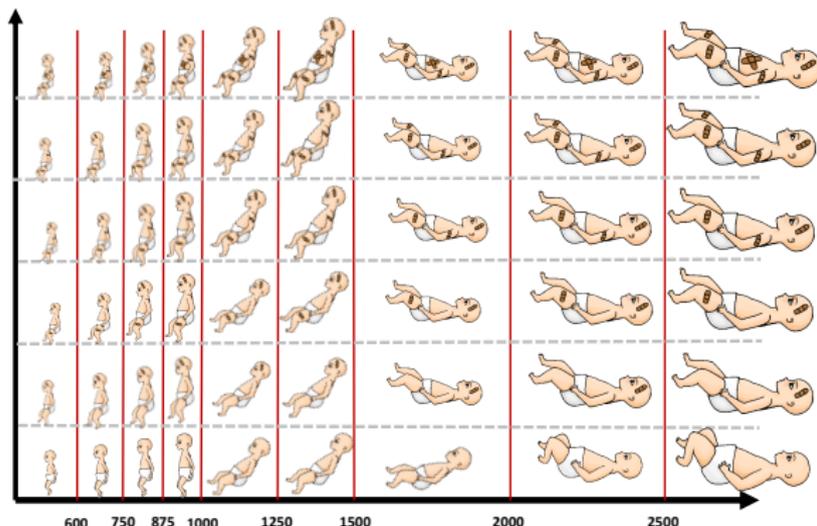
Deutsches Fallpauschalensystem (G-DRG) I

- ▶ Über 1000 Fallgruppen (DRGs), davon 42 DRGs für Neugeborenenbehandlung, davon 37 mit GG.
- ▶ 8 Geburtsgewichtsschwellen (X-Achse) + bis zu 6 Schweregrade (Y-Achse)



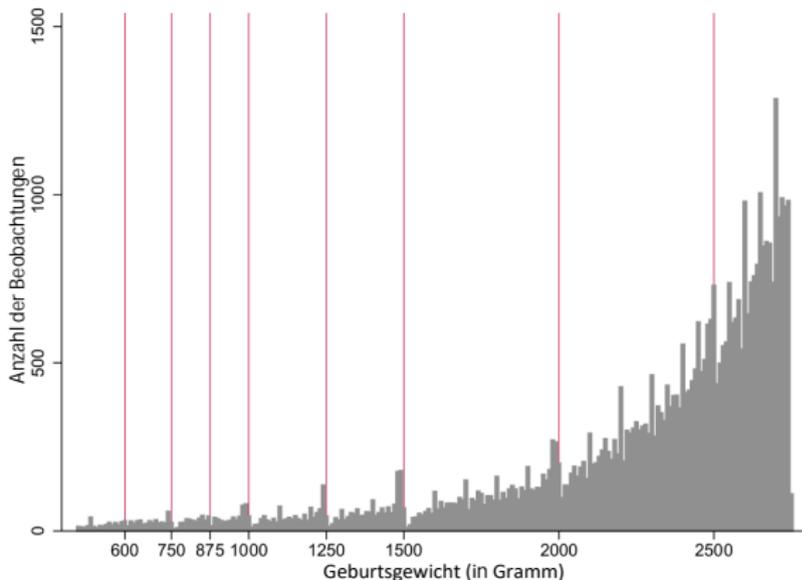
Deutsches Fallpauschalensystem (G-DRG) I

- ▶ Über 1000 Fallgruppen (DRGs), davon 42 DRGs für Neugeborenenbehandlung, davon 37 mit GG.
- ▶ 8 Geburtsgewichtsschwellen (X-Achse) + bis zu 6 Schweregrade (Y-Achse)



Deutsches Fallpauschalensystem (G-DRG) II

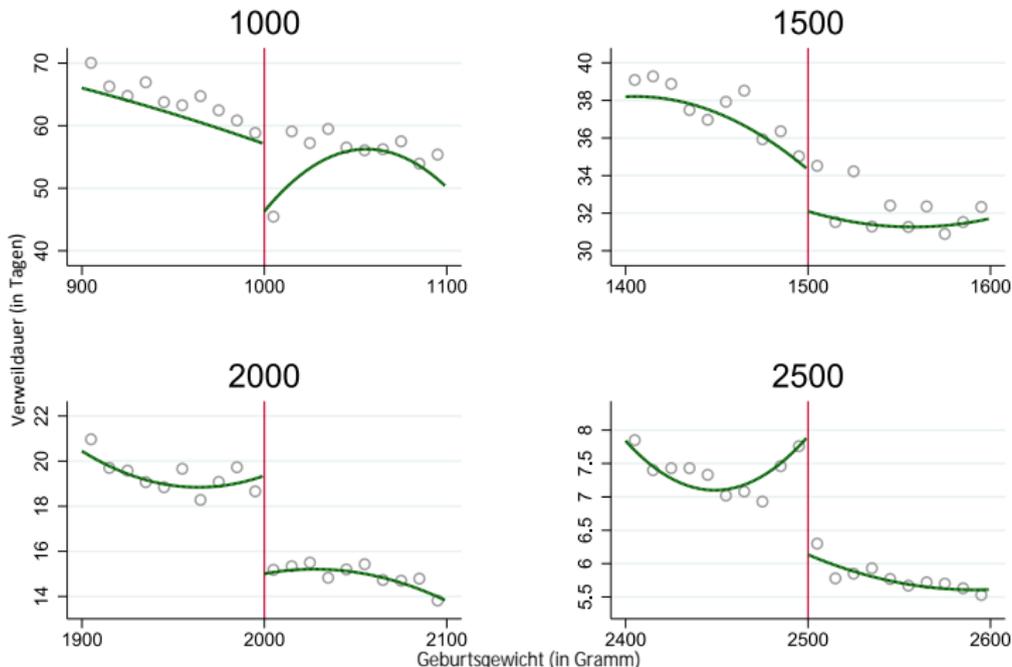
- ▶ Finanzielle Anreize an Geburtsgewichtsschwellen führen zu systematischer Manipulation des berichteten Geburtsgewichts (nach unten).



Datengrundlage

- ▶ Fallpauschalenbezogene Krankenhausstatistik (DRG-Statistik)
- ▶ Datenfernverarbeitung über das FDZ des Bundes und der Länder
- ▶ Gesetzliche Übermittlungsverpflichtung der Krankenhäuser für DRG-Daten an INEK
- ▶ ca. 1 Mio. Geburten (unter 3000g)
- ▶ Vielfältige Informationen: Sozio-demographische Merkmale (Geschlecht, Kreis, GG,...), medizinische Variablen (Fallpauschalen (DRGs), Diagnosen (ICD-10), Prozeduren (OPS), Verweildauer, Verlegungen, Mortalität), Krankenhausvariablen (Abteilungsart)
- ▶ Ergebnisvariablen: Verweildauer, # Prozeduren, Mortalität

Versorgungsintensität - Verweildauer



- ▶ Höhere Verweildauer unterhalb der Schwellen
- ▶ Zudem: Mehr Prozeduren und geringere Mortalität unterhalb der meisten Schwellen, aber ...

Anreize für Upcoding

- ▶ Vergleich von Neugeborenen in einem ± 25 Gram-Intervall um die DRG-Geburtsgewichtsschwellen.
- ▶ Mehr kränkere Neugeborene unterhalb der Schwellen \rightarrow Systematische Unterschiede, die man bei der statistischen Analyse berücksichtigen muss.

	T = 600	T = 750	T = 875	T = 1000	T = 1250	T = 1500	T = 2000	T = 2500
Männlich	.005 (.028)	-.039 (.028)	-.01 (.026)	.021 (.029)	-.003 (.022)	.029 (.023)	.021** (.01)	.031*** (.005)
Ext. Frühgeburt	.014 (.028)	.032 (.028)	.041 (.026)	.085*** (.025)	.01 (.011)	-.006 (.007)	0 (.002)	0 (.001)
Frühgeburt	.002 (.022)	.021 (.023)	-.003 (.024)	.012 (.029)	-.006 (.022)	.024 (.023)	.099*** (.01)	.033*** (.005)
Zwillinge	-.004 (.021)	-.015 (.022)	-.047** (.022)	-.024 (.024)	-.007 (.019)	.003 (.02)	.001 (.01)	.031*** (.004)
Moderate Asphyxie	.011 (.015)	.024* (.014)	-.012 (.013)	-.003 (.014)	.004 (.011)	.018** (.009)	.003 (.004)	.009*** (.001)
IRDS	.045* (.025)	.061*** (.023)	.006 (.021)	.097*** (.027)	.024 (.021)	.119*** (.023)	.111*** (.009)	.081*** (.003)
Mutter raucht	-.006 (.008)	.005 (.007)	-.001 (.009)	.009 (.009)	.003 (.007)	.018*** (.006)	-.001 (.004)	.007*** (.001)
N_U	587	1105	880	1881	1957	3881	5665	14432
N_O	667	446	629	347	693	538	3806	19384

Anmerkungen: Differenz der Mittelwerte unter- und oberhalb der GG-Schwellen (innerhalb einer ± 25 g Bandbreite) ($\bar{X}_U - \bar{X}_O$). Standardfehler unter Mittelwertdifferenz in Klammern. N_U, N_O : Anzahl der Observationen unterhalb bzw. oberhalb der Schwelle innerhalb der 25g Bandbreite. Weitere Angaben zu Kontrollvariablen(-definitionen) im Forschungspapier. *** $p < 0.01$; ** $p < 0.05$ and * $p < 0.1$. **Quelle:** Eigene Berechnungen basierend auf der DRG-Statistik 2005-2011.

Regressionsmethodik

Wir schätzen Differenzen in der Quantität und Qualität der Versorgung als $\hat{\beta}$ mit der folgenden (parametrischen) Regressionsgleichung:

$$Y_i = \alpha + \beta \mathbb{1}(gg_i < T) + f(gg_i) + \gamma \mathbf{X}_i + u_i \quad (1)$$

- ▶ $\mathbb{1}(gg_i < T)$: Indikatorvariable, die Geburten unterhalb der Geburtsgewichtsschwelle T kennzeichnet.
- ▶ Basisspezifikation nutzt ein GG-Polynom 2.Grades $[f(gg_i)]$ und eine große Anzahl an Kontrollvariablen \mathbf{X}_i .
- ▶ Ergebnisvariablen Y_i :
 - ▶ Indikatoren für die quantitative Versorgung: Verweildauer im Krankenhaus (in Tagen) und Anzahl der durchgeführten Prozeduren
 - ▶ Indikatoren für die qualitative Versorgung: Mortalität im Krankenhaus

Regressionsresultate

Bei Berücksichtigung der systematischen (Gesundheits-)Unterschiede der Neugeborenen auf beiden Seiten der GG-Schwellen verschwinden die (meisten) Behandlungsunterschiede.

	T = 600	T = 750	T = 875	T = 1000	T = 1250	T = 1500	T = 2000	T = 2500
Verweildauer (in Tagen)								
(1)	7.963 (5.141)	13.542*** (3.9)	2.055 (2.986)	11.249*** (2.757)	2.604 (1.893)	2.266 (1.745)	4.334*** (.415)	1.733*** (.126)
(4)	-9.439* (5.389)	4.206 (3.567)	-.078 (2.465)	-.457 (2.156)	-.883 (1.746)	-3.446* (1.774)	1.451*** (.367)	.834*** (.113)
Prozeduren (#)								
(1)	1.407 (1.006)	1.427* (.758)	.647 (.656)	1.959*** (.512)	1.049*** (.363)	.246 (.456)	1.122*** (.112)	.859*** (.039)
(4)	-.596 (.964)	-.221 (.707)	.44 (.591)	-.121 (.475)	.508 (.339)	-.569 (.455)	.354*** (.105)	.397*** (.036)
Mortalität								
(1)	-.086** (.041)	-.134*** (.031)	.008 (.022)	-.042* (.023)	-.032** (.014)	-.059*** (.015)	-.011*** (.003)	-.002** (.001)
(4)	.024 (.04)	-.031 (.026)	0 (.018)	.017 (.014)	.002 (.008)	-.001 (.007)	-.005** (.002)	0 (.001)

Anmerkungen: Parametrische Regressionen innerhalb eines $\pm 100g$ Intervalls um die GG-Schwellen. Koeffizient einer binären Variable, die Beobachtungen unterhalb der Schwelle kennzeichnet. (Robuste) Standardfehler in Klammern. Alle Regressionen beinhalten ein GG-Polynom 2. Grades (separat auf beiden Seiten der Schwelle angepasst). Kontrollvariablen: Geschlecht, (extreme) Frühgeburt, Mehrlingsgeburt, Asphyxie, IRDS, Belegabteilung/-arzt/-hebamme, Rauchen, Alkoholismus und Drogenmissbrauch der Mutter, und Jahre. Spezifikation (1) beinhaltet nur den binären GG-Indikator. Spezifikation (4) fügt (1) Kontrollvariablen und Krankenhaus-fixe Effekte hinzu und schränkt die Beobachtungen auf Neugeborene ein, die den ersten Lebensstag überleben. Genaue Variablendefinitionen im Forschungspapier. *** $p < 0.01$; ** $p < 0.05$ and * $p < 0.1$. **Quelle:** Eigene Berechnungen basierend auf der DRG-Statistik 2005-2011

Sensitivitätsanalysen

- ▶ Placebo-Schwellen
- ▶ Kontrolle für Schweregrad-Gruppen
- ▶ Treatment-Effekt-Bounds
- ▶ Bonferoni-Korrektur für multiples Hypothesen-Testen
- ▶ Ausschluss aller Geburten, die vor dem 4. Lebenstag sterben
- ▶ Analysen für einzelne Jahre
- ▶ "Donut"-Regressionen
- ▶ Verschiedene funktionale Formen der laufenden Variable (GG)

Fazit

- ▶ Frage: Beeinflussen jetzt GG-Schwellen die Gesundheitsversorgung?
 - ▶ Antwort: Kommt drauf an!
 - ▶ Ärzte geben ein geringeres GG für Neugeborene mit mehr Gesundheitsproblemen an.
 - ▶ Diese (kränkeren) Neugeborenen erhalten mehr Gesundheitsversorgung.
 - ▶ Wenn man aber vergleichbare Neugeborene auf beiden Seiten der Geburtsgewichtsschwellen vergleicht, gibt es wenig Evidenz dafür, dass diese unterschiedlich behandelt werden.
 - ▶ Das spricht dafür, dass Ärzte im Interesse ihrer Patienten handeln und das DRG-System nur nutzen, um auf Fallbasis kein "Minusgeschäft" zu machen.
- ▶ Haben wir jetzt alle Fragen zu diesem Thema beantwortet (können)?
 - ▶ Gute Nachricht: Nein → Neue interessante Forschungsprojekte möglich!
 - ▶ Schlechte Nachricht: Nein → Einige wichtige Fragen bei bisheriger Datenlage nicht beantwortbar!

Wünsche, Vorschläge, etc.

Zitat: "Deutschland soll zum weltweit führenden Standort für KI werden. [...] **Die Menge an nutzbaren qualitativ hochwertigen Daten muss deutlich erhöht werden**, ohne dabei Persönlichkeitsrechte, das Recht auf informationelle Selbstbestimmung oder andere Grundrechte zu verletzen." Quelle: Entwurf Eckpunktepapier der Bundesregierung

Verbesserung des Datenzugangs, des Datenumfangs, der Linkage-Möglichkeiten, der Kontinuität, ... → Viele, aufwendige Forderungen, aber rechtfertigt der Aufwand den Ertrag? **JA! Warum?**

- ▶ Standortfaktor
- ▶ Öffentliches Gut → Steigerung des Wohlfahrt durch mehr Forschung
- ▶ Private Daten (KVs, Ärzte, ...) meist kein Substitut
- ▶ Big Data lieber beim Stat. Landesamt als bei Google

Wünsche, Vorschläge, etc.

Konkrete Wünsche für die DRG-Statistik:

- ▶ Trägerschaft des KHS (Privat vs. Öffentlich vs. Frei-gemeinnützig)
- ▶ Versicherungsstatus (Privat vs. Gesetzlich)
- ▶ Verknüpfung von Müttern und Neugeborenen
- ▶ Verknüpfung über die Zeit (verschiedene Hospitalisierungsepisoden)
- ▶ Direkte, interaktive Datenfernverarbeitung
- ▶ ...

Ein großer Dank an die Mitarbeiter des FDZs des Stat. Bundesamts (Nutzerbetreuung, Fachabteilung) und des FDZ des Bay. Stat. Landesamts!

Danke für die Aufmerksamkeit!
Kommentare, Kritik, oder Vorschläge?!
Wichert@ifo.de