

"Wichtig ist nicht, ob ein Resultat sexy ist, sondern ob die Fragestellung relevant ist"

Auf Initiative des Thurgauer Wirtschaftsinstituts (TWI) wird das Fachjournal "Experimental Economics" in einer Spezialausgabe Forschungsarbeiten veröffentlichen, die es bereits vor dem Vorliegen der Resultate begutachtet hat. Der Entscheid für die Publikation hängt damit von der Bedeutung der Fragestellung und des vorgestellten Forschungsdesigns ab und nicht vom Ergebnis. Ziel ist, zu zeigen, dass auch die Veröffentlichung von Null-Resultaten für den Wissenschaftsbetrieb gewinnbringend ist. Damit will das TWI einen Beitrag zur Diskussion um die Replikationskrise leisten – also um die Tatsache, dass sich manche Resultate etwa in der psychologischen, medizinischen oder auch ökonomischen Forschung nicht replizieren lassen. Im Interview erklärten Prof. Dr. Urs Fischbacher, Leiter des TWI, und Irenaeus Wolff, Post-Doc-Researcher am Institut, die Beweggründe zu ihrer Initiative.

Über die Publikation einer wissenschaftlichen Arbeit zu entscheiden, noch bevor das Resultat vorliegt, klingt zunächst einmal widersinnig. Was soll das der Wissenschaft bringen?

I. Wolff: Das hat eine ganze Reihe von Vorteilen. Am wichtigsten ist, dass es das Problem des Publikations-Bias löst. In einigen Wissenschaften, auch in den Wirtschaftswissenschaften, werden nämlich hauptsächlich Studien publiziert, die einen vermuteten Zusammenhang zeigen. Solche, die ihn nicht zeigen, werden eher nicht zur Publikation angenommen oder gar nicht erst aufgeschrieben.

U. Fischbacher: Genau, und dabei sind diese sogenannten «Null-Resultate» immens wichtig, um ein adäquates Gesamtbild zu erhalten. Ein eng damit verbundener Punkt ist, dass wir wieder anfangen müssen, die Fragestellung in den Vordergrund zu stellen. Wichtig ist, ob diese relevant ist und nicht, ob ein Effekt «sexy» genug ist, um Aufmerksamkeit zu generieren. Solange wir nur Arbeiten begutachten, deren Ergebnisse wir kennen, wird es in der Regel darum gehen, Artikel zu publizieren, die möglichst viel Aufmerksamkeit finden. Wissenschaftlerinnen und Wissenschaftler sind halt auch nur Menschen, die gerne beachtet werden.

I. Wolff: Und schliesslich gibt es noch einen ganz praktischen dynamischen Effekt: stellen Sie sich mal vor, 20 Leute haben dieselbe (falsche) Vermutung. Dann wäre es sinnvoll, wenn die ersten fünf, die die Vermutung erfolglos testen, ihre Ergebnisse veröffentlichen – dann können sich die restlichen 15 die Mühe und das Geld sparen, dem gleichen vermeintlichen Effekt hinterherzurennen.

Konkret: Wieso soll die Publikation eines Null-Resultats einen Erkenntnisgewinn bringen?

U. Fischbacher: Weil es zum Beispiel auch wichtig ist zu wissen, welche Medikamente nicht gegen Krebs helfen. Völlig klar aber wird es, wenn wir den ganzen Weg anschauen, der zu einer Publikation führt: Nehmen wir einmal an, weltweit dächten 100 Wissenschaftler, dass Störche die Kinder bringen. Nun erheben alle 100 dazu Daten. Bei einem "Signifikanzniveau" von fünf Prozent werden rein zufällig etwa fünf einen Zusammenhang zwischen dem Storchbestand und der Geburtenzahl feststellen. Wer aber schreibt die Artikel dazu? Natürlich die fünf, die den spannenden Effekt beobachtet haben, während die andern, die keinen Effekt sehen, sich denken: "Das kann ich sowieso nicht publizieren", und ihre Resultate noch nicht einmal zu Papier bringen. Und wenn erstmal die ersten Studien "mit Effekt" dazu veröffentlicht sind, dann wird es für die an-

deren noch schwieriger, da dann meist von ihnen verlangt wird zu erklären, warum in ihrer Studie der "wissenschaftlich belegte" Effekt nicht auftritt.

I. Wolff: Wenn aber nun die Studien bereits vor Erhebung der Daten begutachtet werden, so passieren zwei Dinge: es werden mehr "Null-Resultate" veröffentlicht. Das kommt der Realität näher. Umgekehrt macht es auch die Resultate, also die beobachteten Effekte, zuverlässiger, weil pure Zufälle eher erkannt werden. Wir erhalten somit in beiden Fällen bessere Informationen: Wenn es keinen Zusammenhang gibt, weil mehr Information sichtbar wird, und wenn es einen Zusammenhang gibt, weil die sichtbare Information verlässlicher ist.

Eigentlich müssten publizierte falsche Resultate eines Wissenschaftlers durch seine Kolleginnen und Kollegen falsifiziert werden. Funktioniert das Falsifizierungsprinzip denn nicht mehr?

U. Fischbacher: Das Problem liegt vor allem darin, dass die Interpretation einer Falsifikation nicht immer eindeutig ist. Es kann daran liegen, dass die erste Studie zufällig ein extremes Ergebnis erzielt hat, das sich dann in der zweiten Studie nicht repliziert. Es kann aber auch sein, dass die zweite Studie etwas anders durchgeführt wurde, dass sie etwas „falsch“ gemacht hat. Deshalb sind solche Falsifikationen nicht leicht zu publizieren. Letztlich wäre es aber am besten, wenn die erste Studie und viele weitere Replikationen publiziert würden. Man erhielte so ein immer besseres Bild über den genauen Effekt. Einerseits über die Grösse des Effekts, aber auch darüber, unter welchen Bedingungen sich der Effekt einstellt.

I. Wolff: Das Problem ist also, dass eine Nicht-Replikation selten eindeutig ist. Das hat auch damit zu tun, dass wir es in der Wirtschaftswissenschaft eigentlich nie mit mechanischen und daher deterministischen Effekten zu tun haben. Vor diesem Hintergrund sind also sowohl die Erststudie wie auch die Replikation letztendlich "Zufallsstichproben". Man hat also nach einer Nicht-Replikation eine Stichprobe, die auf einen Effekt hinweist, und eine, die darauf hinweist, dass es den Effekt nicht gibt. Das kann aber sowohl dann passieren, wenn es den Effekt gibt, als auch, wenn es den Effekt nicht gibt.

U. Fischbacher: Somit sind sowohl Erststudie als auch "Falsifikation" zwar wichtig und wertvoll, die einzelne Nicht-Replikation erlaubt aber in der Regel noch kein abschliessendes Urteil über das (Nicht-)Vorhandensein des untersuchten Effekts. Das macht sie für Zeitschriften und Autorinnen und Autoren natürlich nicht so attraktiv wie die zeitgleich eingereichten, brandaktuellen nächsten 20 Studien, die alle einen anderen neuen Effekt zweifelsfrei zu belegen scheinen.

Auch für ein wissenschaftliches Fachmagazin dürfte es attraktiver sein, Artikel mit aufsehenerregenden Erkenntnissen zu publizieren statt Null-Resultate. Wie konnten sie "Experimental Economics" davon überzeugen, das in einer Spezialausgabe das Risiko von Null-Resultaten einzugehen?

U. Fischbacher: Diese Spezialausgabe soll uns helfen, Fragen zu beantworten, deren Beantwortung interessant sind, und zwar unabhängig von der Antwort. Wir wollen so von der Sichtweise wegkommen, dass Null-Resultate riskant sind, denn Null-Resultate enthalten, wie bereits ausgeführt, eine wichtige Information.

I. Wolff: Letzten Endes hat wohl die Einsicht gesiegt, dass die Vor-Resultate-Begutachtung viele Probleme löst und kaum neue schafft, und schliesslich ist es auch erst ein Versuch. Wir werden sehen, wie es funktioniert.

U. Fischbacher: Es ist ja eigentlich auch im Interesse der Fachmagazine, dass Artikel, die bei ihnen erscheinen, auch wirklich hieb- und stichfest sind. Gerade die sogenannte

Replikationskrise sollte uns das lehren. Die Tatsache, dass sich eine grosse Anzahl von teilweise Aufsehen erregenden Experimenten nicht replizieren liessen, hat das Vertrauen in wissenschaftliche Studien doch empfindlich gestört.

"Experimental Economics" lässt sich jedenfalls auf den Versuch ein. Bis Januar 2020 können Wissenschaftlerinnen und Wissenschaftler Arbeiten zur Vor-Resultate-Begutachtung einreichen. Welche Rolle kommt dem TWI im Begutachtungsprozess zu?

I. Wolff: Wir werden als «Gasteditoren» dafür zuständig sein, geeignete Experten als Gutachter zu gewinnen und letztendlich aufgrund von deren Empfehlungen die Entscheidungen darüber zu treffen, welche Projekte im Zuge des Pilotprojekts realisiert werden sollen.

Mit wie vielen eingereichten Arbeiten rechnen Sie?

I. Wolff: Das ist schwierig einzuschätzen. Die einzige wirtschaftswissenschaftliche Zeitschrift, die bislang ein solches Unterfangen gestartet hat, hatte bis Dezember 2018 21 Einreichungen. Wir hoffen auf eine mindestens ähnlich grosse Resonanz. Wir haben bereits einige sehr positive Rückmeldungen, teilweise auch klare Absichtserklärungen, etwas einreichen zu wollen. Wir haben Platz für fünf bis sieben Artikel, die dann in der Herbst- oder Winterausgabe von Experimental Economics erscheinen werden.

Glauben Sie, dass Sie mit der Spezialausgabe in "Experimental Economics" etwas bewegen können?

U. Fischbacher: Das hoffen wir natürlich. Tatsächlich hinken wir Ökonomen in diesem Bereich deutlich hinter der Psychologie her – dort gibt es schon über hundert Zeitschriften, die eine Vor-Resultate-Begutachtung anbieten.

I. Wolff: Aber auch bei uns scheint sich die Einsicht langsam durchzusetzen, dass die Vor-Resultate-Begutachtung eine sinnvolle Sache ist. Es gibt eine ganze Reihe von prominenten Forscherinnen und Forschern, die sich dafür ausgesprochen und uns zu unserer Initiative beglückwünscht haben.



Prof. Dr. Urs Fischbacher



Dr. Irenaeus Wolff