

Quasi-experimentelle Ansätze in der quantitativen Wirkungsforschung zur Sozialen Arbeit im städtischen Raum

Andreas Bengesser

Abstract Der Beitrag behandelt die Anwendung quasi-experimenteller Methoden zur quantitativen Wirkungsmessung in drei Bereichen: die Auswirkungen der Mobilen Jugendarbeit und der Gemeinwesenarbeit auf strafrechtlich relevante Handlungen sowie auf das subjektive Sicherheitsempfinden und die objektive Sicherheit im städtischen öffentlichen Raum. Unter Verwendung umfangreicher Daten, die von der öffentlichen Verwaltung in Österreich erfasst und aufbereitet wurden, werden die kausalen Effekte dieser sozialen Projekte evaluiert. Fortgeschrittene statistische Methoden wie Panel-Regression, Propensity-Score-Matching und Vergleiche mit synthetischen Kontrollgruppen werden eingesetzt, um zuverlässige kausale Schlussfolgerungen zu ermöglichen. Diese Methoden werden im Beitrag ausführlich behandelt. Es wird illustriert, wie robuste methodische Ansätze erfolgreich in der sozialarbeiterischen Praxis implementiert werden können und betont die Bedeutung interdisziplinärer Zusammenarbeit in der Stadtforschung.

Keywords Wirkungsmessung; quasi-experimentelle Methoden; Fixed-Effect-Panelregression; Propensity-Score-Matching; synthetische Kontrollgruppen

Messung kausaler Effekte in der sozialräumlichen Sozialarbeit

Interdisziplinäre Ansätze in der Stadtforschung zielen u.a. darauf ab, ein ganzheitliches Bild der Wechselwirkungen zwischen (sozialen) Interventionen und städtischem Raum zu zeichnen. Durch die Verknüpfung von Methoden und theoretischen Perspektiven aus Disziplinen wie Soziologie, Psychologie, Ökonomie und Soziale Arbeit, ermöglicht dieser Ansatz das präzise Messen und Verstehen der oft »verborgenen« Effekte Sozialer Arbeit im städtischen öffentlichen Raum.

Um diese »verborgenen« Effekte sichtbar zu machen, ist die Messung und Bewertung der Effektivität sozialer Interventionen im städtischen öffentlichen Raum von zentraler Bedeutung. Dies stellt jedoch oft eine Herausforderung dar. Evalua-

tion und Wirkungsmessung sind methodische Ansätze, die genutzt werden, um Interventionen systematisch zu analysieren und ihre Effekte auf die Gesellschaft zu verstehen. Evaluation beschreibt den systematischen Prozess der Datenerhebung, -analyse und -interpretation, um die Wertigkeit, Zielerreichung, Effizienz und Nachhaltigkeit von Interventionen zu beurteilen. Wirkungsmessung hingegen fokussiert spezifisch auf die quantitative Bewertung der Veränderungen, die durch bestimmte Interventionen hervorgerufen werden, und strebt an, kausale Effekte isoliert darzustellen. Diese Methoden sind fundamental für evidenzbasierte Entscheidungen und die Optimierung des Ressourceneinsatzes in verschiedenen sozialen und politischen Bereichen. Beispielsweise kann eine Wirkungsmessung dazu beitragen, die Effektivität von Programmen zur Gewaltprävention in benachteiligten Stadtteilen zu bewerten.

Verschiedene statistische Methoden können den Zusammenhang zwischen Ereignissen zeigen. Die Frage nach einem kausalen Effekt – einer Veränderung allein aufgrund einer sozialen Intervention – ist jedoch komplexer und mit verschiedenen Herausforderungen verbunden. Forschende stoßen auf Probleme wie die Identifikation von Störfaktoren oder Konfundierungsvariablen¹, die den beobachteten Zusammenhang beeinflussen können. Auch die Auswahl der Stichprobe und die Art der Datenerhebung können erheblichen Einfluss auf die Validität der Schlussfolgerungen haben. Daher erfordert die Forschung nicht nur geeignete statistische Methoden, sondern auch eine sorgfältige Berücksichtigung und Kontrolle dieser spezifischen Probleme, um zuverlässige Aussagen über kausale Effekte treffen zu können.

Prinzipien der quasi-experimentellen Forschung

Um die Herausforderungen der Kausalitätsforschung zu bewältigen, haben sich experimentelle Ansätze als gewinnbringend herausgestellt. In einem Experiment werden die Teilnehmenden zufällig einer Interventionsgruppe (die die Maßnahme erhält) oder einer Kontrollgruppe (die keine Maßnahme erhält) zugeordnet. Durch diese Randomisierung wird sichergestellt, dass sowohl bekannte als auch unbekannte Störfaktoren gleichmäßig auf beide Gruppen verteilt sind. Dies ermöglicht den Forschenden, einen kausalen Effekt der Intervention zu isolieren und zu messen. Ein Beispiel hierzu aus der Medizin: Angenommen ein Pharmaunternehmen möchte die Wirksamkeit eines neuen Medikaments zur Behandlung einer bestimmten Krankheit überprüfen. Die Forschenden würden Patient:innen zufällig entweder der Gruppe, die das neue Medikament erhält, oder der Kontrollgruppe,

1 Konfundierung bezieht sich auf eine Situation in der Statistik, in der die Beziehung zwischen zwei Variablen verfälscht oder irreführend erscheint, weil eine Drittvariable (*Konfundierungsvariable*) mit beiden in Verbindung steht.

die ein Placebo bekommt, zuweisen. Durch diese Randomisierung wird sichergestellt, dass mögliche Einflussfaktoren wie Alter, Geschlecht oder Schweregrad der Erkrankung gleichmäßig auf beide Gruppen verteilt sind.

Diese Vorgehensweise stößt jedoch auf Grenzen, wenn es um Untersuchung von Nutzer:innen städtischer öffentlicher Räume geht. In einem städtischen Umfeld ist es nicht möglich, Menschen zufällig verschiedenen Umgebungen zuzuweisen, wie es in einem kontrollierten Experiment gefordert wäre. Stattdessen müssen Forschende alternative Methoden, wie quasi-experimentelle Designs, verwenden. In solchen Fällen kann die Identifizierung von Kausaleffekten herausfordernder sein, da es schwieriger ist, alle Störfaktoren zu kontrollieren, die das Verhalten der Menschen beeinflussen könnten. Für die sozialräumliche Soziale Arbeit sind diese Erkenntnisse besonders relevant, da sie sich auch mit der Gestaltung und Bewertung von Interventionen in öffentlichen und gemeinschaftlichen städtischen Räumen beschäftigt. Sozialarbeiter:innen und Planer:innen können diese Forschungsergebnisse nutzen, um gezielte Maßnahmen zu entwickeln, die die Lebensqualität und das soziale Miteinander in städtischen Gebieten verbessern. Sie müssen jedoch kreative und methodisch robuste Ansätze entwickeln, um die Effekte dieser Maßnahmen in der komplexen Dynamik städtischer Umgebungen zu erfassen und zu verstehen.

Ziel der im Beitrag vorgestellten Methoden ist es, ein experimentelles Design zu replizieren. In Situationen, in denen eine echte Randomisierung nicht möglich oder aus ethischen Gründen nicht vertretbar ist, bietet die quasi-experimentelle Forschung eine alternative Herangehensweise. Ein zentrales Merkmal quasi-experimenteller Zugänge ist ebenfalls die Nutzung von Kontrollgruppen, teilweise auch in Verbindung mit einem Vorher-Nachher-Vergleich. Anders als bei experimentellen Studien mit randomisierten Interventions- oder Kontrollgruppen greifen quasi-experimentelle Studien auf bereits existierende Gruppen zurück. In diesem Kontext wird versucht, die Äquivalenz zwischen Interventions- und Kontrollgruppen durch verschiedene Methoden herzustellen. Dabei spielen Kontrollgruppen eine entscheidende Rolle, um Störfaktoren zu minimieren und eine genauere Bewertung der Effekte von Interventionen zu ermöglichen. Der Vorher-Nachher-Vergleich innerhalb der Gruppen erlaubt es, Veränderungen im zeitlichen Verlauf zu erfassen und die spezifischen Auswirkungen der Intervention genauer zu analysieren. Dieses Vorgehen ermöglicht damit eine differenzierte Untersuchung von Ursache-Wirkungs-Zusammenhängen, auch wenn eine vollständige Randomisierung nicht durchführbar ist.

Methode 1: Paneldaten – mit unbeobachteten Effekten durch gezieltes Fixieren umgehen

Paneldaten beziehen sich auf Daten, bei denen dieselben Einheiten wie Individuen, Regionen oder Unternehmen usw. über mehrere Zeitpunkte hinweg beobachtet werden. Dadurch ermöglichen diese Daten eine Analyse von Veränderungen über die Zeit auf individueller Ebene und bieten somit eine dynamische Perspektive für die Untersuchung von Entwicklungen, Trends und kausalen Zusammenhängen. Paneldaten werden oft mithilfe von Panelregressionen analysiert, einer statistischen Methode, die es ermöglicht, sowohl individuelle als auch zeitliche Variationen in den Daten zu berücksichtigen. Es gibt verschiedene Arten von Panelregressionen, darunter Fixed-Effect-Panelregressionen, Random-Effect-Panelregressionen und Pooled-Panelregressionen, die jeweils unterschiedliche Ansätze zur Modellierung von individuellen Effekten und zeitlichen Veränderungen bieten. Im Folgenden liegt der Fokus auf Fixed-Effect-Panelregressionen, da diese Methode einen detaillierten Blick auf die Effekte innerhalb der beobachteten Einheiten ermöglicht. Diese Fokussierung erlaubt eine präzise Analyse der individuellen Veränderungen über die Zeit, was für Wirkungsanalysen von besonderer Relevanz ist.

Fixed-Effect-Panelregressionen

Fixed-Effect-Panelregressionen zielen darauf ab, individuelle Effekte (Fixed Effects) oder unveränderliche Unterschiede zwischen den Beobachtungseinheiten zu berücksichtigen. Diese Methode wird angewendet, um potenzielle Störgrößen zu kontrollieren², die über die Zeit konstant bleiben und somit nicht durch herkömmliche Regressionsmodelle, wie das OLS-Modell (Ordinary Least Squares), erfasst werden. Dafür wird für jede individuelle Beobachtungseinheit eine Dummy-Variable eingeführt, die individuellen Unterschiede repräsentiert. Diese Dummy-Variable nimmt den Wert 1 für die spezifische Beobachtungseinheit und den betrachteten Zeitraum an sowie 0 für alle anderen. Durch die Einbeziehung dieser Fixeffekte wird die Modellierung von individuellen Unterschieden ermöglicht,

-
- 2 In Panelregressionen bezieht sich das ›Kontrollieren‹ darauf, bestimmte Variablen in das Modell aufzunehmen, um potenzielle Störeinflüsse zu berücksichtigen und die Genauigkeit der geschätzten Effekte zu verbessern. Kontrollvariablen werden eingeführt, wenn die beobachteten Effekte zwischen der unabhängigen Variable (derjenigen, die untersucht wird) und der abhängigen Variable (derjenigen, die gemessen wird) nicht auf andere Faktoren zurückzuführen sein sollen. Kontrollvariablen dienen dazu, unbeabsichtigte Variationen oder Drittvariablen zu isolieren und die Robustheit der geschätzten Beziehung zwischen den Hauptvariablen zu gewährleisten. Insgesamt ermöglicht das Kontrollieren in Panelregressionen eine präzisere Analyse von Zusammenhängen über die Zeit und trägt dazu bei, alternative Erklärungen für beobachtete Effekte auszuschließen.

während gleichzeitig die zeitliche Varianz innerhalb der Beobachtungseinheiten analysiert werden kann. Die Fixed-Effect-Panelregression ermöglicht somit die Kontrolle von beobachtungseinheitenspezifischen unbeobachteten oder beobachtbaren Faktoren und bietet eine robuste Methode zur Untersuchung von zeitlichen Veränderungen innerhalb des untersuchten Panels (Wooldridge 2019).

Paneldaten zeichnen sich durch wiederholte Messungen von einer oder mehreren Variablen – wie Deliktzahlen, Sicherheitsbewertungen, Umfragen zur Wohnzufriedenheit, schulische Leistungen, Teilnahme an außerschulischen Aktivitäten usw. – in einer oder mehreren Regionen aus. Im Vergleich zu Querschnittsdaten bieten sie mehrere Vorteile:

- Sie ermöglichen die Analyse individueller Entwicklungen, da sowohl die Varianz zwischen den Regionen als auch die Varianzen innerhalb der Regionen berücksichtigt werden.
- Sie enthalten Informationen über die zeitliche Abfolge von Zuständen (z.B. Deliktzahlen) und Ereignissen (z.B. Interventionen durch die Soziale Arbeit).
- Sie erlauben die Kontrolle unbeobachteter individueller (regionaler) Heterogenität.
- Sie sind informativer als Querschnittsdaten, da sie mehr Variabilität und Freiheitsgrade aufweisen.
- Sie ermöglichen einen Vorher-Nachher-Vergleich sowie einen Vergleich mit Kontrollgruppen bzw. Kontrollregionen.

Fixed-Effect-Modelle in der Panelregression fokussieren auf die Untersuchung von Veränderungen innerhalb derselben Einheiten (zum Beispiel Stadtteile oder Personen) über die Zeit. Diese Methode ist besonders vorteilhaft, wenn das Ziel darin besteht, die Effekte von Variablen zu isolieren, die sich innerhalb der Einheiten im Zeitverlauf verändern, während alle unveränderlichen Charakteristika der Einheiten kontrolliert werden. Durch die Fokussierung auf die Intra-Unit-Veränderungen können einige fehlende Variablen kontrolliert werden, insbesondere wenn diese konstant über die Zeit sind. Externe Einflüsse, die sich im Laufe der Zeit ändern, könnten in Fixed-Effect-Modellen immer noch eine Rolle spielen, jedoch wird der Einfluss von zeitinvarianten Variablen durch die Annahme eliminiert, dass diese konstant bleiben und somit nicht die abhängige Variable beeinflussen.

Anwendungsbeispiel: Evaluation Mobile Jugendarbeit

Die Studie³ zur Wirkungsevaluation Mobiler Jugendarbeit (Bengesser 2017) veranschaulicht die Relevanz der Fixed-Effect-Panelregression in der Evaluierung sozialräumlicher Interventionen. Die Mobile Jugendarbeit richtet sich an sozial besonders vulnerable Jugendliche und junge Erwachsene mit prekären gesellschaftlichen Inklusionschancen, die oft als soziale Risikogruppen wahrgenommen werden. Diesen wird im öffentlichen Diskurs ein erhöhtes Risiko für kriminelle Handlungen oder ein delinquenter Lebensweg zugeschrieben. Die Interventionen der Mobilen Jugendarbeit zielen auf individueller und kollektiver Ebene darauf ab, die Entwicklung dieser Zielgruppe positiv zu beeinflussen, um Gewalt und Kriminalität vorzubeugen. Die Untersuchung hatte zum Ziel, die Auswirkungen dieser sicherheitsrelevanten Intervention auf strafrechtlich relevante Handlungen im geografischen Einflussbereich der Mobilen Jugendarbeit im Großraum Wien und im östlichen Niederösterreich zu untersuchen. Die Analyse stützt sich auf Daten des österreichischen Sicherheitsmonitors (SIMO), einem landesweiten Datenanalyse- und Speicherungssystem, das von Exekutivbeamt:innen zur Aufklärung und Prävention von Straftaten genutzt wird (Lattacher 2004). Insgesamt standen Daten von vierzehn Gebieten – acht Interventionsgebiete Mobiler Jugendarbeit und sechs Vergleichsgebiete – zwischen Jänner 2005 und August 2014 zur Verfügung. Diese Paneldaten liefern präzise Informationen über Art, Zeitpunkt und Ort der Delikte. Um die Untersuchung auf Delikte der Zielgruppe von Jugendarbeitsinterventionen zu fokussieren, wurden nur solche berücksichtigt, bei denen der Polizei tatverdächtige Personen bekannt waren. Die Delikte wurden entsprechend den vorrangigen Zielgruppen der Mobilen Jugendarbeit ausgewählt. Die Daten ermöglichen einen Vorher-Nachher-Vergleich sowie einen Vergleich mit den sechs Kontrollregionen, in denen die Mobile Jugendarbeit nicht tätig war. Die räumliche Auswahl der Delikte orientierte sich am Aktionsradius verschiedener Einrichtungen der Mobilen Jugendarbeit.

Um den kausalen Effekt der Mobilen Jugendarbeit zu ermitteln, ist es notwendig regionsspezifische Einflussfaktoren zu berücksichtigen, die sowohl die Inan-

3 Das von Jänner 2014 bis Mai 2016 unter Projektleitung des IRKS – Institut für Rechts- und Kriminalsoziologie durchgeführte KIRAS-Forschungsprojekt JA_SICHER – »Jugendarbeit im öffentlichen Raum als mehrdimensionale Sicherheitsmaßnahme: Ansätze zur Wirkungsevaluation« verfolgte die Zielsetzungen, wissenschaftlich fundiertes Wissen über die Wirkungen dieser sicherheitsrelevanten Interventionsmaßnahmen zu gewinnen. Gemeinsam mit dem FORSAR – Forschungszentrum für Soziale Arbeit der FH Campus Wien und in enger Kooperation mit zwei Praxispartnern der Offenen Jugendarbeit in Wien (Verein Wiener Jugendzentren) und Niederösterreich (Verein TENDER) sowie dem Bundesministerium für Inneres als öffentlichem Bedarfsträger wurden Wirkungsindikatoren identifiziert, gegenstandsadäquate Forschungsinstrumente zur Wirkungsevaluation erarbeitet und an Jugendarbeitsprojekten in Wien und Niederösterreich angewandt.

spruchnahme Mobiler Jugendarbeit als auch das Auftreten strafrechtlich relevanter Handlungen beeinflussen können. Zu diesen Faktoren zählen sozioökonomische Bedingungen wie die Arbeitslosenquote, das durchschnittliche Einkommensniveau und die Verfügbarkeit von Bildungsangeboten. Ebenfalls von Bedeutung sind lokale politische Entwicklungen, wie Veränderungen in der Sicherheitspolitik und eine erhöhte Polizeipräsenz. Zudem sind spezifische Merkmale der Zielgruppen entscheidend, die bei der Interpretation der Ergebnisse eine wichtige Rolle spielen. Diese vielschichtigen Einflussgrößen müssen sorgfältig analysiert werden, um die tatsächlichen Auswirkungen der Mobilen Jugendarbeit präzise bewerten zu können. Da die vorhandenen Paneldaten mehr als zwei Zeitpunkte umschließen und Fixed-Effect-Panelregressionen auf mehr als zwei Zeitpunkte Analysen liefern, wurde diese Methode angewandt. Die Auswertung erfolgte in vier Fixed-Effect-Panelmodellen⁴, die verschiedene Variablen wie Interventionen, Interventionsdauer, Trendwende, saisonale Unterschiede und Periodeneffekte berücksichtigten:

1. Modell 1 konzentriert sich primär auf den Einfluss der Interventionen der Mobilen Jugendarbeit auf Straftaten und vernachlässigt dabei saisonale Unterschiede, die u. a. auf Witterungseinflüsse zurückzuführen sind. Dadurch erhält man eine grundlegende Einsicht in den direkten Zusammenhang zwischen Interventionen und Kriminalität.
2. Modell 2 erweitert Modell 1, indem es zusätzlich zu den Interventionen auch saisonale Unterschiede, insbesondere Quartals-Effekte, berücksichtigt. Diese Erweiterung ermöglicht eine genauere Analyse, da die saisonalen Schwankungen auch unabhängig von den Interventionen auftreten können.
3. Modell 3 geht noch einen Schritt weiter und integriert neben den Variablen aus Modell 1 und 2 auch regionsunspezifische Periodeneffekte. Dies ist besonders relevant, wenn sich zu einem bestimmten Zeitpunkt systematische Bedingungen in allen Regionen ändern, beispielsweise bei der Datenerfassung. Durch diese Berücksichtigung wird die Modellierung robuster gegenüber möglichen systematischen Veränderungen.
4. Modell 4 erweitert die Analyse durch die Einbeziehung der Einsatzdauer der Mobilen Jugendarbeit. Es geht der Frage nach, ob eine längere Einsatzdauer zu einem verstärkten Rückgang der Delikte führt oder ob die Wirkung mit der Zeit nachlässt. Diese Erweiterung ermöglicht eine differenzierte Betrachtung der zeitlichen Dimension und trägt dazu bei, die langfristigen Effekte der Interventionen besser zu verstehen, insbesondere in verschiedenen Regionen.

4 In Panelregressionen bezieht sich das ›Modell‹ auf die mathematische Darstellung der Beziehung zwischen den untersuchten (abhängigen) Variablen und unabhängigen (erklärenden) Variablen sowie möglichen Kontrollvariablen.

Aussagen, Beurteilung und Kritik

Zusammenfassend kann gesagt werden, dass die schrittweise Erweiterung der Modelle dazu dient, immer mehr relevante Aspekte zu berücksichtigen. Dies ermöglicht eine umfassendere Analyse der Zusammenhänge zwischen den Interventionen der Mobilen Jugendarbeit und den Straftaten, wodurch präzisere und signifikante Ergebnisse erzielt werden können. Die Ergebnisse der Studie zeigten überraschende Zusammenhänge, die auf eine Verringerung von Straftaten mit jugendlichen Tatverdächtigen im Einflussbereich der Mobilen Jugendarbeit hindeuteten. Diese Ergebnisse, obwohl statistisch abgesichert, bleiben schwer interpretierbar aufgrund von unbekannten Kontextfaktoren. Daher werden sie lediglich als starke Wirkungsindizien betrachtet. Zudem ermöglichen Fixed-Effect-Panelregressionen keine Einblicke in die Wirkungsmechanismen der Mobilen Jugendarbeit. Es bleibt unklar, ob die Reduktion von Straftaten auf Verhaltensänderungen der Jugendlichen oder auf andere Faktoren, wie das Ausweichen in andere Gebiete, eine erhöhte Toleranz im sozialräumlichen Umfeld oder alternative Konfliktlösungen, zurückzuführen sind. Um diese Fragen zu klären, könnten sogenannte White-Box-Verfahren (Ottmann/König 2023: 134ff), die sowohl quantitative als auch qualitative Ansätze nutzen, hilfreich sein.

Obwohl die Studie Ansätze eines Vergleichsgruppendesigns verfolgt, wurden bestimmte Einschränkungen in Kauf genommen. Die Auswahl der Vergleichsregionen berücksichtigte zwar soziodemografische Kriterien und wurde von den Fachkräften der Mobilen Jugendarbeit als »vergleichbar« beurteilt. Dennoch lässt sich nicht ausschließen, dass bei der Auswahl eine gewisse Subjektivität eine Rolle spielte. Daher werden in den nachfolgenden Kapiteln Methoden präsentiert, die vielversprechend erscheinen, um dieses Problem der Implementierung eines validen Kontrollgruppendesigns zu bewältigen.

Methode 2: Propensity-Score-Matching

Matching-Verfahren stellen eine Methode dar, um in nicht-experimentellen Situationen, nachträglich Vergleichsgruppen zu erstellen, ähnlich der Randomisierung in Experimenten. Dabei wird z. B. jeder Person, jedem Stadtteil usw. die oder der von einer sozialräumlichen Maßnahme profitiert hat (Behandlungsgruppe), eine vergleichbare Person bzw. ein vergleichbarer Stadtteil usw. zugeordnet, die oder der nicht von der Maßnahme betroffen war (Kontrollgruppe). Das Ziel besteht darin, dass die Eigenschaften der beiden Gruppen möglichst identisch sind. Ein oft verwendetes Matching-Verfahren ist das *Propensity-Score-Matching*, bei dem Beobachtungen mit ähnlichen Wahrscheinlichkeiten für die Zuweisung einer Behandlung einander zugeordnet werden (Luellen et al. 2005).

Propensity-Score-Matching ist eine statistische Methode, die in nicht-randomisierten Studiendesigns verwendet wird, um Vergleichsgruppen zu erstellen und potenzielle Verzerrungen zu minimieren. Der Propensity-Score repräsentiert die Wahrscheinlichkeit, dass ein Individuum die beobachtbaren Merkmale aufweist. Der Prozess des Propensity-Score-Matching besteht aus mehreren Schritten: Zu Beginn wird der Propensity-Score für jedes Individuum in der Studienpopulation geschätzt. Dies erfolgt in der Regel durch die Anwendung statistischer Modelle wie der logistischen Regression⁵. Die beobachtbaren Merkmale, die als Prädiktoren⁶ dienen, können demografische Faktoren oder andere relevante Variablen sein. Nach der Schätzung folgt der Matching-Prozess. Individuen in der Interventionsgruppe werden mit solchen in der Kontrollgruppe paarweise abgeglichen, basierend auf ähnlichen oder identischen Propensity-Scores. Hierbei können verschiedene Matching-Techniken angewendet werden, darunter exaktes Matching, Nearest-Neighbor-Matching oder Kernel-Matching (Steiner/Cook 2013:237ff). Das Ziel des Matchings besteht darin, vergleichbare Gruppen zu bilden, sodass die Einflüsse von potenziell konfundierenden Variablen⁷ minimiert werden. Indem Personen mit ähnlichen Propensity-Scores miteinander verglichen werden, wird die Wahrscheinlichkeit erhöht, dass beobachtbare Merkmale zwischen den Gruppen ausgewogen sind. Nach dem Matching können die Ergebnisse der Interventions- und Kontrollgruppe verglichen werden. Dies ermöglicht eine robustere Analyse der kausalen Effekte, da potenzielle Verzerrungen durch unterschiedliche Hintergrundmerkmale minimiert werden. Propensity-Score-Matching unterstützt somit die Bildung von vergleichbaren Gruppen in nicht-randomisierten Studien und trägt dazu bei, die interne Validität zu stärken, indem es die Auswirkungen von Konfundierung reduziert (Rosenbaum/Rubin 1983).

Zusammenfassend lässt sich die quasi-experimentelle Methode des Propensity Scoring wie folgt charakterisieren:

-
- 5 Die logistische Regression ist eine statistische Methode, um die Beziehung zwischen einer binären abhängigen Variable (z. B. Ja/Nein) und mehreren unabhängigen Variablen zu analysieren. Im Unterschied zur linearen Regression, die für kontinuierliche abhängige Variablen geeignet ist, kommt die logistische Regression bei kategorialen Variablen zum Einsatz. Sie modelliert die Wahrscheinlichkeit des Eintretens des Ereignisses (z.B. Erfolg) als Funktion der unabhängigen Variablen.
 - 6 In einer logistischen Regression bezieht sich der Begriff »Prädiktoren« auf die unabhängigen Variablen, welche die Wahrscheinlichkeit eines Ereignisses vorherzusagen.
 - 7 Konfundierende Variablen sind sowohl mit der unabhängigen als auch mit der abhängigen Variable verbunden. Wenn sie nicht kontrolliert werden, können sie die Studienergebnisse verzerren. Die Kontrolle stellt sicher, dass der beobachtete Effekt auf die unabhängige Variable zurückzuführen ist und nicht auf andere Einflussfaktoren.

Propensity Scoring ist eine statistische Methode, die in der Sozialwissenschaft, Epidemiologie und Ökonometrie verwendet wird, um den Einfluss von Konfundierungsvariablen zu minimieren und kausale Effekte genauer zu schätzen.

Hintergrund: Propensity Scoring basiert auf der Annahme, dass die Zuordnung zu einer Behandlungsgruppe nicht zufällig ist, sondern von bestimmten Faktoren beeinflusst wird.

Propensity Score: Der Propensity Score ist die Wahrscheinlichkeit eine bestimmte Behandlung basierend auf beobachtbaren Kovariablen zu erhalten. Er wird durch Anwendung von logistischen Regressionsmodellen oder anderen Methoden geschätzt.

Matching oder Gewichtung: Es werden verschiedene Methoden angewendet, um behandelte und nicht-behandelte Gruppen mit ähnlichen oder gewichteten Propensity Scores zu vergleichen. Matching bedeutet, dass für jeden Fall ein ähnlicher nicht-behandelter Fall gefunden wird. Die Gewichtung gibt unterschiedliche Gewichtungen für behandelte und nicht-behandelte Gruppen basierend auf ihren Propensity Scores, vor.

Reduzierung der Konfundierung: Durch das Angleichen versucht man, Konfundierungsvariablen zu kontrollieren und somit die Validität der kausalen Schlussfolgerungen zu verbessern.

Herausforderungen: Die Qualität der Propensity Scores ist abhängig von der Genauigkeit der Modellschätzung, wobei potenzielle Fallstricke, wie beispielsweise nicht beobachtbare Konfundierungsvariablen, existieren können.

Anwendungsbeispiel: Subjektive Sicherheit im städtischen öffentlichen Raum

Ein Beispiel für die Anwendung des Propensity-Score-Ansatzes findet sich in der Auswertung der standardisierten Befragung zur Raumnutzung und subjektiven Sicherheit im Rahmen des Projekts »community work's« (Bengesser 2023a). Dieses KIRAS-Forschungsprojekt⁸ setzt sich mit den Auswirkungen von Gemeinwesenarbeit (GWA) im öffentlichen Raum, insbesondere im Kontext der Sicherheit, auseinander. Zwischen 2020 und 2022 wurden verschiedene Ansätze der GWA in Österreich mittels eines multimethodischen Forschungsdesigns untersucht. Dieses Design kombinierte qualitative sozialräumliche Fallstudien in sechs unterschiedlichen Standorten mit standardisierten Face-to-Face-Befragungen im öffentlichen Raum.

8 Das österreichische Sicherheitsforschungsprogramm KIRAS ist ein nationales Programm der Österreichische Forschungsförderungsgesellschaft (FFG) zur Förderung der Sicherheitsforschung in Österreich.

Die Auswahl der Standorte erfolgte nach theoretischen Sampling-Kriterien, um eine Vielfalt an sozio-ökonomischen und kulturellen Kontexten sowie GWA-Ansätzen abzudecken⁹. Die qualitative Forschung basierte auf ethnographischen Methoden (Rogojanu/Wolfmayr 2024 in diesem Band) wie (teilnehmender) Beobachtung und Expert:inneninterviews. Parallel dazu wurden standardisierte Befragungen zur subjektiven Sicherheit im öffentlichen Raum durchgeführt. Diese Befragung zielte auch darauf ab, die sicherheitsrelevanten Auswirkungen der Anwesenheit von GWA-Mitarbeiter:innen im öffentlichen Raum zu erfassen. Es zeigt sich, dass den Befragten überwiegend (insgesamt: 52 %) nicht auffiel, dass Sozialarbeiter:innen oder ähnliche Berufsgruppen im untersuchten Sozialraum aktiv waren.

Um die Auswirkungen von GWA auf das Sicherheitsempfinden der Befragten im konkreten Sozialraum präzise zu bewerten, genügt es nicht, lediglich einen Vergleich des durchschnittlichen Sicherheitsempfindens beider Gruppen durchzuführen – also derjenigen, die die GWA-Fachkräfte wahrnahmen, und derjenigen, die sie nicht wahrnahmen. Neben soziodemografischen Merkmalen können auch andere Faktoren, wie das allgemeine Sicherheitsempfinden in der Stadt oder am Ort, spezifische wahrgenommene Raummerkmale, negative Erlebnisse oder die Häufigkeit des Aufenthalts am Platz, das Sicherheitsempfinden beeinflussen. Um die Wirkungen von GWA möglichst präzise zu erfassen, ist es entscheidend, potenzielle Einflussfaktoren zu berücksichtigen. Mit dem Ziel, ausschließlich kausale Zusammenhänge bezüglich der Auswirkungen der Maßnahmen zu untersuchen, wurden Vergleichsgruppen mithilfe des Propensity-Score-Matchings bestimmt (Rosenbaum/Rubin 1983). In der Studie wurden nur Personen paarweise verglichen, die sich in der Wahrnehmung der GWA-Fachkräfte unterschieden, aber hinsichtlich der Wahrscheinlichkeit, die GWA-Fachkräfte wahrnehmen zu können, nicht unterschiedlich waren. Die Analyse des Effekts einer wahrgenommenen Anwesenheit der Fachkräfte erfolgte anschließend durch multiple Regressionsanalysen¹⁰.

Wenn nicht nur Mitarbeiter:innen der untersuchten Projekte, sondern auch solche der GWA bzw. Sozialen Arbeit, die den Projekten nicht direkt zugeordnet wer-

9 Es wurden insgesamt fünf GWA-Einrichtungen ausgewählt, die sich durch ihre intensive Tätigkeit im öffentlichen Raum auszeichnen. Die Auswahl umfasst Einrichtungen in verschiedenen geografischen Kontexten: eine Großstadt (Wien), eine Mittelstadt (Salzburg) und kleinere bzw. ländliche Orte (Tirol). Diese Einrichtungen repräsentieren unterschiedliche ethnische und sozioökonomische Sozialräume. Sie variieren in ihrer Ausrichtung zwischen ›bottom-up‹ und ›top-down‹, staatlicher und zivilgesellschaftlicher Verankerung sowie zwischen emanzipatorischen und wohlfahrtsstaatlichen, tendenziell konsensorientierten Ansätzen. Dabei berücksichtigen sie Sicherheitsaspekte sowohl implizit als auch explizit.

10 Multiple Regressionsanalysen dienen dazu die Beziehung zwischen einer abhängigen Variablen (hier: das subjektive Sicherheitsgefühl) und mehreren unabhängigen Variablen zu untersuchen.

den konnten, berücksichtigt wurden, zeigte sich ein positiver, statistisch signifikanter Effekt auf das Sicherheitsempfinden von Besucher:innen der fokussierten öffentlichen Räume. Im Durchschnitt stieg das Sicherheitsempfinden bei Befragten, die Fachkräfte der GWA bzw. der Sozialen Arbeit im betreffenden öffentlichen Raum wahrnahmen, um etwa einen Bewertungspunkt auf der zehnstufigen Skala (Bengesser 2023a).

Aussagen, Beurteilung und Kritik

Die Ergebnisse deuten darauf hin, dass sich die Anwesenheit von Fachkräften der GWA-Projekte als auch anderer Einrichtungen positiv auf das Sicherheitsempfinden von Besucher:innen von Parks bzw. öffentlichen Räumen auswirkt. Die qualitativen Daten der Untersuchung zeigen darüber hinaus, dass aufsuchende Soziale Arbeit, insbesondere mit sozialräumlicher Ausrichtung, die subjektive Sicherheit marginalisierter Menschen im öffentlichen Raum positiv beeinflussen kann. In urbanen Räumen mit Irritationen und Widersprüchen kann die »urbane Gelassenheit« (Fritsche et al. 2024) durch GWA gesteigert und Prozesse des Ausschlusses sowie Verdrängung reduziert werden. Konflikte können durch von GWA angeleitete Prozesse produktiv und inklusiv bearbeitet und die Beteiligung artikulationsschwacher Gruppen an demokratischen Prozessen, insbesondere bei der Gestaltung öffentlicher Räume, kann erhöht werden. Räumliche Gestaltungen fördern Kommunikations- und Bildungsprozesse. GWA trägt zur Bewältigung generalisierter Ängste in Bezug auf gesellschaftliche Transformationsprozesse bei und ermöglicht Selbstwirksamkeitserfahrungen, indem abstrakte Sicherheitsbedenken direkt angegangen werden können.

Diese Ergebnisse dienen auch als Beispiel für die jeweiligen Stärken der methodischen Zugänge (Van-Hametner/Smigiel 2024 in diesem Band): Während quantitative Methoden unter bestimmten Voraussetzungen feststellen können, ob eine Wirkung nachweisbar ist, können qualitative Methoden analysieren, warum diese eingetreten ist.

Methode 3: Synthetische Kontrollgruppen-Methode

Während beim Propensity-Score-Matchings reale Kontrollgruppen zur Verfügung stehen, ermöglicht die Synthetische Kontrollgruppen-Methode eine präzise Untersuchung in Situationen, in denen keine tatsächliche Kontrollgruppe verfügbar ist.

Die synthetische Kontrollgruppen-Methode ist daher eine weitere statistische Technik, um kausale Effekte von Interventionen oder Ereignissen auf Gesamteinheiten wie Städte, Regionen oder Länder zu analysieren. Sie bietet eine Antwort auf die Herausforderungen herkömmlicher vergleichender Fallstudien, bei denen

die Auswahl der Kontrollgruppen oft subjektiv erfolgt und Bedenken hinsichtlich der Repräsentativität aufwirft. Diese Methode wurde von Abadie und Gardeazabal (2003) sowie Abadie et al. (2010) entwickelt und zielt darauf ab, eine »synthetische Kontrolleinheit« zu erstellen, die eine Art »künstliche« Vergleichsgruppe für eine behandelte Einheit bildet. Dies ist besonders nützlich in Studien ohne ideale, reale Kontrollgruppe.

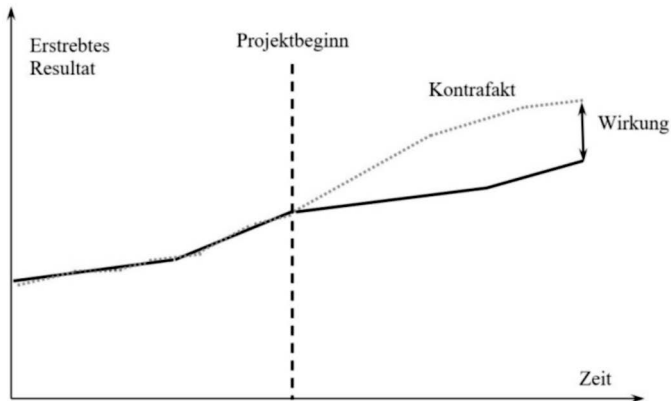
Die Anwendung der synthetischen Kontrollgruppen-Methode erfolgt typischerweise in drei Schritten.

- (1) Zunächst werden eine oder mehrere behandelte Einheiten identifiziert, die einer Intervention ausgesetzt waren. Dann wird ein Pool von potenziellen Kontrolleinheiten festgelegt, die als Basis für die synthetische Kontrolleinheit dienen. Dies sind Einheiten, die nicht der Intervention ausgesetzt waren, aber ähnliche Merkmale wie die behandelte Einheit vor der Intervention aufweisen.
- (2) Im nächsten Schritt werden Gewichtungen für jede Kontrolleinheit im Pool berechnet, um eine synthetische Kontrolleinheit zu erstellen. Die Berechnung dieser Gewichtungen ist ein entscheidender Schritt, um sicherzustellen, dass die synthetische Kontrolleinheit die relevanten Merkmale der behandelten Einheit vor der Intervention am besten widerspiegelt. Die Gewichtungen basieren auf einem Optimierungsprozess, der darauf abzielt, die Ähnlichkeit zwischen der behandelten Einheit und der synthetischen Kontrolleinheit vor Interventionsbeginn zu maximieren¹¹. Die Gewichtungen geben somit an, wie stark jede Kontrolleinheit zur Bildung der synthetischen Kontrolleinheit beiträgt. Eine höhere Gewichtung bedeutet, dass die entsprechende Kontrolleinheit einen größeren Einfluss auf die synthetische Kontrolleinheit hat. Durch diesen Anpassungsprozess wird sichergestellt, dass die synthetische Kontrolleinheit eine repräsentative Nachbildung der behandelten Einheit darstellt, indem sie eine ausgewogene Kombination der verfügbaren Kontrolleinheiten berücksichtigt.
- (3) Schließlich werden die Auswirkungen der Intervention auf die behandelte Einheit durch den Vergleich der beobachteten Ergebnisse mit den synthetisch geschätzten Ergebnissen ermittelt (siehe Abb. 1). Dieser Vergleich erfolgt durch einen sorgfältigen Abgleich der tatsächlich beobachteten Ergebnisse für die behandelte Einheit (durchgehende Linie in Abb. 1.) mit den Ergebnissen, die

11 Ein typisches Vorgehen besteht darin, die Gewichtungen so anzupassen, dass die synthetische Kontrolleinheit eine möglichst geringe Differenz zu den Merkmalen der behandelten Einheit vor Interventionsbeginn aufweist. Dies kann mithilfe von mathematischen Optimierungsalgorithmen, wie beispielsweise dem »Least Squares Matching« erreicht werden. Ziel ist es, eine Kombination von Gewichtungen zu finden, die die Summe der quadratischen Unterschiede zwischen den Merkmalen der behandelten Einheit und der synthetischen Kontrolleinheit vor Interventionsbeginn minimiert.

aus der synthetischen Kontrolleinheit abgeleitet wurden (gestrichelte Linie in Abb. 1). Dabei werden Abweichungsdiagramme verwendet, die die Differenz zwischen den Ergebnissen der behandelten Einheit und der synthetischen Kontrolle über die Zeit visuell darstellen.

Abb. 1: Schematische Darstellung Kontrafakt



Quelle: eigene Darstellung

Abbildung 1 veranschaulicht die grundlegende Herausforderung der Wirkungsmessung, bei der der Versuch unternommen wird, den Unterschied zwischen dem tatsächlich eingetretenen Ergebnis (durchgehende Linie) und dem Kontrafakt (gepunktete Linie) zu quantifizieren.

Für die Überprüfung der Signifikanz einer Intervention werden Placebo-Tests verwendet. Bei diesen Tests wird jede der unbehandelten Einheiten so behandelt, als ob sie die Intervention zur gleichen Zeit erhalten hätte wie die tatsächlich behandelte Einheit. Dies dient dazu, eine Vergleichsbasis zu schaffen und die Signifikanz des beobachteten Effekts der tatsächlichen Behandlung zu überprüfen. Durch die Simulation von Placebo-Behandlungen an den Kontrollgruppen lässt sich eine Verteilung der Placebo-Effekte erstellen. Die beobachteten Effekte der tatsächlich behandelten Einheit werden dann mit dieser Verteilung verglichen, um die statistische Signifikanz zu bewerten. Um die Robustheit der Ergebnisse sicherzustellen, können verschiedene Robustheitstests¹² durchgeführt werden. Ein wichtiger Robustheits-

12 Robustheitstests in diesem Kontext beziehen sich darauf, die Stabilität und Zuverlässigkeit der Ergebnisse nach einem Vergleich mit synthetischen Kontrollgruppen sicherzustellen. Diese Tests prüfen, ob die geschätzten Effekte oder Unterschiede zwischen der behandelten Gruppe und der synthetischen Kontrollgruppe konsistent und widerstandsfähig gegen-

test bezieht sich auf die künstliche Variation der Startzeiten der Intervention, um zu untersuchen, ob die geschätzten Effekte stabil bleiben oder wie sie sich verändern.

Die *Synthetische Kontrollgruppen-Methode* ist eine Ergänzung zu traditionellen experimentellen Ansätzen. Sie ermöglicht eine kausale Wirkungsschätzung in vergleichenden Studien, auch wenn keine echte Kontrollgruppe vorhanden ist. Sie basiert auf der Idee, eine synthetische Vergleichsgruppe zu schaffen, die ähnliche Eigenschaften wie die behandelte Gruppe aufweist. Diese Vergleichsgruppe wird aus einer Kombination von Vergleichseinheiten gebildet, die ähnliche Eigenschaften wie die behandelte Gruppe aufweisen.

Anwendungsbeispiel: Objektive Sicherheit im städtischen öffentlichen Raum

Ein konkretes Beispiel aus dem Bereich der sozialräumlichen Sozialen Arbeit ist abermals das Projekt »community work's« (Bengesser 2023b). In diesem Projekt wurde auch untersucht, ob während der Aktivitätszeiträume der betreffenden Projekte Veränderungen bei den von der Polizei erfassten Straftaten beobachtet werden konnten. Zur Identifikation adäquater Vergleichsregionen, die nicht von der Maßnahme beeinflusst waren, wurde die synthetische Kontrollgruppen-Methode angewendet. Dafür wurde die Zusammensetzung der realen Vergleichsregionen¹³ so ausgewählt, dass ihre Summe möglichst nah an die Anzahl der polizeilich erfassten Straftaten vor Beginn der Projekte in der betrachteten Region herankommt. Anders ausgedrückt wurden die Kontrollregionen nicht willkürlich, sondern datengesteuert zusammengesetzt. Dies ermöglicht eine Annäherung an die Frage, wie sich die polizeilich erfassten Straftaten entwickelt hätten, wenn die untersuchten Projekte nicht aktiv gewesen wären.

über verschiedenen Modellspezifikationen sind. Durch Robustheitstests wird die Gültigkeit der Methode überprüft und gewährleistet, dass die Ergebnisse nicht stark von spezifischen Modellannahmen oder Parametereinstellungen abhängen.

- 13 Bei der Auswahl der realen Vergleichsregionen werden demografische, sozioökonomische, geografische und kriminologische Merkmale berücksichtigt. Beispiele hierfür sind Bevölkerungsdichte, Einkommensniveau, Bildungsniveau, Arbeitslosenquote und historische Kriminalitätsrate. Um Größenunterschiede zu kompensieren, wird eine Normalisierung der Daten vorgenommen. Dafür werden die Straftaten pro 1.000 Einwohner berechnet. Wenn die Kriminalitätsrate stark von der Größe des Gebiets abhängt, können die Daten auch pro Quadratkilometer berechnet werden.

Aussagen, Beurteilung und Kritik

Im Unterschied zu den oben genannten Ergebnissen der standardisierten Befragung weichen in der SIMO-Analyse die Entwicklungen der Regionen mit GWA-Interventionen von den entsprechenden Entwicklungen in den jeweiligen synthetischen Kontrollregionen statistisch nicht signifikant ab. D.h., einerseits zeigen die Fragebogendaten eine Auswirkung der GWA-Präsenz auf das subjektive Sicherheitsempfinden, andererseits sind Veränderungen der objektiven Sicherheitslage auf Basis der SIMO-Daten statistisch nicht nachweisbar. Inwiefern andere, z.B. nicht von der Statistik erfasste bzw. nicht angezeigte oder nicht ins Blickfeld der Exekutive gelangte Straftaten beeinflusst werden, bleibt offen.

Obwohl diese Ergebnisse die Auswirkungen von GWA auf die objektive Sicherheitslage in Frage stellen, bestätigen sie dennoch den Einfluss auf das subjektive Sicherheitsempfinden. Diese Befunde verdeutlichen jedoch nicht, auf welche Weise GWA wirkt. Sie zeigen stattdessen auf, dass die Beziehungen zwischen objektiver und subjektiver Unsicherheit komplex und nicht zwingend kausal determiniert sind (Hirtenlehner/Hummelsbacher 2015: 459ff.). Ein tiefgehendes Verständnis der spezifischen, komplexen Einflüsse von GWA auf öffentliche städtische Räume wird erst durch die Analyse der qualitativen Daten ermöglicht.

Conclusio

Die Analyse urbaner Prozesse und der Auswirkungen städtischer Maßnahmen mit Hilfe quasi-experimenteller Techniken unterstreicht die Wichtigkeit eines fachübergreifenden Zugangs in der Stadtforschung. Durch das Zusammenführen von Wissen aus Feldern wie der Sozialwissenschaften, Verhaltenswissenschaften, Wirtschaftswissenschaften und der Sozialen Arbeit gelingt es, ein umfassendes Bild der vielschichtigen Dynamiken in Stadtgebieten zu zeichnen. Mit Methoden wie der Fixed-Effects-Panelregression, dem Propensity-Score-Matching und den synthetischen Kontrollgruppen erlangen Forscher:innen profunde Einblicke in die kausalen Wirkungen sozialer Projekte und Interventionen (Angrist/Pischke 2010; Imbens/Wooldridge 2009).

Diese Methoden, ursprünglich aus der Ökonometrie¹⁴ stammend, erscheinen auf den ersten Blick nicht direkt interdisziplinär. Ihr interdisziplinärer Charakter entsteht jedoch durch den Kontext und die Art und Weise, wie sie eingesetzt werden, um komplexe Fragestellungen in urbanen Räumen zu untersuchen, die mehrere Disziplinen berühren. Beispielsweise analysiert die Fixed-Effect-Panelregression Daten über Zeit und Individuen hinweg, um interne Effekte zu isolieren und langfristige soziale sowie ökonomische Auswirkungen politischer Maßnahmen zu verstehen. Propensity-Score-Matching ermöglicht es, Vergleichsgruppen in Nicht-Experimentalsituationen zu schaffen, wobei Soziolog:innen und Psycholog:innen sicherstellen, dass die Gruppenbildung präzise ist und relevante Variablen berücksichtigt. Die Methode der synthetischen Kontrollgruppen zieht Daten aus diversen Quellen heran, um eine künstliche Vergleichsgruppe zu schaffen, die realen Kontexten gegenübergestellt wird. Dabei werden Erkenntnisse aus Statistik, Wirtschaftswissenschaften und Stadtplanung genutzt.

Die interdisziplinäre Zusammenarbeit ist entscheidend für die effektive Anwendung dieser Methoden, da sie die Vielfalt städtischer Lebenswelten erfasst und die Effekte von Interventionen auf eine Weise analysiert, die ohne die Kombination von Kenntnissen aus verschiedenen Bereichen nicht möglich wäre. Dies unterstreicht die Notwendigkeit in der Stadtforschung disziplinübergreifende Teams zu bilden, die in der Lage sind, die komplexe Realität urbaner Interventionen zu entschlüsseln und zu verstehen. Diese Teams tragen nicht nur zu präziseren Forschungsergebnissen bei, sondern verbessern auch die Planung und Durchführung von Stadtpolitik und -management auf eine informierte und evidenzbasierte Weise.

Ein wesentlicher Aspekt der Wirkungsforschung ist die Fähigkeit, nicht nur zu bestimmen, ob bestimmte Maßnahmen wirken, sondern auch zu verstehen, warum und unter welchen Bedingungen sie Erfolg haben oder scheitern. Dieses tiefere Verständnis der Mechanismen hinter den Effekten hängt oft von der Kombination quantitativer und qualitativer Forschungsmethoden ab. Während quantitative Methoden robuste Daten liefern, um die Wirksamkeit von Interventionen zu messen, ermöglichen qualitative Methoden Einblicke in die Erfahrungen und Wahrnehmungen der Betroffenen.

Die praktische Relevanz der Forschungsergebnisse zeigt sich besonders in der Möglichkeit, politische Entscheidungsträger:innen und Praktiker:innen mit soliden

14 Ökonometrie ist ein Teilgebiet der Wirtschaftswissenschaften, das mathematische und statistische Methoden verwendet, um ökonomische Daten zu analysieren und Theorien zu testen. Sie dient der quantitativen Analyse ökonomischer Phänomene und unterstützt die Formulierung von wirtschaftspolitischen Entscheidungen durch die Schaffung ökonometrischer Modelle. Diese Modelle helfen, ökonomische Beziehungen zu quantifizieren, Vorhersagen über zukünftige Trends zu treffen und die Wirkung ökonomischer Politiken zu bewerten.

Daten zu versorgen, die eine fundierte Grundlage für die Gestaltung städtischer Interventionen bieten. Darüber hinaus fordert die städtische Dynamik eine kontinuierliche Überprüfung und Anpassung der Maßnahmen. Die adaptiven und flexiblen Forschungsansätze sind daher von zentraler Bedeutung, um auf neue Herausforderungen und Veränderungen in städtischen Umgebungen effektiv reagieren zu können.

Abschließend sei hervorgehoben, dass zukünftige Forschungen in der interdisziplinären Stadtforschung weiterhin innovative methodische Ansätze entwickeln sollten, um die Komplexität urbaner Interventionen besser zu erfassen. Die Integration von fortschrittlichen Technologien wie Big Data und künstlicher Intelligenz könnte dabei helfen, Echtzeitanalysen städtischer Phänomene zu ermöglichen und so die Grundlage für proaktive und präventive städtische Strategien zu stärken. Durch das Zusammenführen der verschiedenen Perspektiven und Methoden kann die Stadtforschung weiterhin einen entscheidenden Beitrag zur Verbesserung der Lebensqualität in urbanen Zentren leisten und gleichzeitig dazu beitragen, die Herausforderungen moderner Städte nachhaltig zu adressieren.

Literatur

- Abadie, Alberto/Gardeazabal, Javier (2003): *Economic Costs of Conflict. A Case Study of the Basque Country*, in: American Economic Review 93(1), 113–132.
- Abadie, Alberto/Diamond, Alexis/Hainmueller, Jens (2010): *Synthetic Control Methods for Comparative Case Studies. Estimating the Effect of California's Tobacco Control Program*, in: Journal of the American Statistical Association 105(490), 493–505.
- Angrist, Joshua D./Pischke, Jörn-Steffen (2010): *The credibility revolution in empirical economics. How better research design is taking the con out of econometrics*, in: Journal of Economic Perspectives 24(2), 3–30.
- Bengesser, Andreas (2023a): Standardisierte Befragung zu Raumnutzung und subjektiver Sicherheit, in: Andrea Fritsche/Hemma Mayrhofer/Andreas Bengesser/Christoph Stoik: *Endbericht zu KIRAS-Projekt community work's – Gemeinwesenarbeit als Sicherheitsfaktor im öffentlichen Raum. Wirkerkenntnisse und Erfolgsfaktoren*. Wien.
- Bengesser, Andreas (2023b): Wirkungsevaluation ausgewählter GWA-Projekte anhand von Daten des österreichischen Sicherheitsmonitors (SIMO), in: Andrea Fritsche/Hemma Mayrhofer/Andreas Bengesser/Christoph Stoik: *Endbericht zu KIRAS-Projekt community work's – Gemeinwesenarbeit als Sicherheitsfaktor im öffentlichen Raum: Wirkerkenntnisse und Erfolgsfaktoren*. Wien.
- Bengesser, Andreas (2017): Wirkungsevaluation mobiler Jugendarbeit anhand von Daten des Sicherheitsmonitors, in: Hemma Mayrhofer (Hg.), *Wirkungsevaluation*

- mobiler Jugendarbeit. Methodische Zugänge und empirische Ergebnisse*. Opladen: Verlag Barbara Budrich, 287–310.
- Fritsche, Andrea/Bengesser, Andreas/Stoik, Christoph/Mayrhofer, Hemma (2024): *Gemeinwesenarbeit als Sicherheitsfaktor im öffentlichen Raum? Ergebnisse der Wirkungsevaluation sozialräumlicher Projekte im Rahmen des KIRAS-Projekts »community work's«*, in: SIAK-Journal, Edition 1/2024, 47–58.
- Hirtenlehner, Helmut/Hummelsheim, Dina (2015): Kriminalitätsfurcht und Sicherheitsempfinden. Die Angst der Bürger vor dem Verbrechen (und dem, was sie dafür halten), in: Nathalie Guzy/Christoph Birkel/Robert Mischkowitz (Hg.), *Viktimisierungsbefragungen in Deutschland*, Wiesbaden: Bundeskriminalamt, 458–487.
- Imbens, Guido W./Wooldridge, Jeffrey M. (2009): *Recent developments in the econometrics of program evaluation*, in: Journal of Economic Literature 47(1), 5–86.
- Lattacher, Siegbert (2004): *Sicherheitsmonitor. Frühwarnsystem bei Kriminalität*, in: Öffentliche Sicherheit 3–4, https://bmi.gv.at/magazinfiles/2004/03_04/files/sicherheitsmonitor.pdf [Zugriff am 20.11.2023].
- Luellen, Jason K./Shadish, William R./Clark, M. H. (2005): *Propensity scores. An introduction and experimental test*, in: Evaluation Review 29(6), 530–558.
- Ottmann, Sebastian/König, Joachim (2023): *Wirkungsorientierung in der Sozialen Arbeit*. Stuttgart: Kohlhammer.
- Rosenbaum, Paul R./Rubin, Donald B. (1983): *The central role of the propensity score in observational studies for causal effects*, in: Biometrika 70, 41–55.
- Steiner, Peter M./Cook, David (2013): Matching and propensity scores, in: Todd D. Little (Ed.), *Oxford handbook of quantitative methods*, Oxford: Oxford University Press, 237–259.
- Wooldridge, Jeffrey (2019): *Introductory Econometrics. A Modern Approach*. 7th Edition, Cengage Learning: South-Western.

