

Norad



World Food Programme



Federal Ministry
for Economic Cooperation
and Development
BMZ



DIME

DATI E VALUTAZIONE PER
MIGLIORARE LA POLITICA
PUBBLICA

Arianna Legovini

alegovini@worldbank.org

IERI

- i dati scarsi e costosi da ottenere
- la ricerca economica è progredita a un ritmo lento, specialmente nei paesi poveri di dati
- Il costo della valutazione d'impatto determinata dagli alti costi dei dati

OGGI

- 2.5 quintilioni di bites al giorno,
- 3.6 miliardi di utenti di smart phones
- L'uso dei dati è in rapida espansione specialmente nel settore privato
- Ma la gran parte dei dati non viene utilizzata, soprattutto nel settore pubblico
- La scarsità adesso è la capacità di ricerca

BIG DATA

Statistics

\$



Un processo sistematico e iterativo per espandere la capacità tecnica attraverso il ciclo di ricerca

Costruire le capacità

Formazione

- [Manage Successful Impact Evaluations](#) (2014-2020)
- [RA Technical Onboarding](#)
- Software trainings: [R](#), [Stata](#), [GitHub](#), [LaTeX](#), [SurveyCTO](#)
- DIME/CEGA/DG [Measuring Development Conference](#) (2015-20)

Open-source Risorse

- [DIME Wiki](#): 128,000 users from 200 countries
- [DIME Analytics Data Handbook](#)
- Open-source tools: [ietoolkit](#), [iefieldkit](#), [data visual library](#)

Garantire la qualità

Revisione tecnica del progetto di ricerca e dei risultati

Revisione tecnica esterna in doppio cieco/revisione interna/autorizzazione etica/registrazione di prova/controllo della versione

Garanzia della qualità dei dati

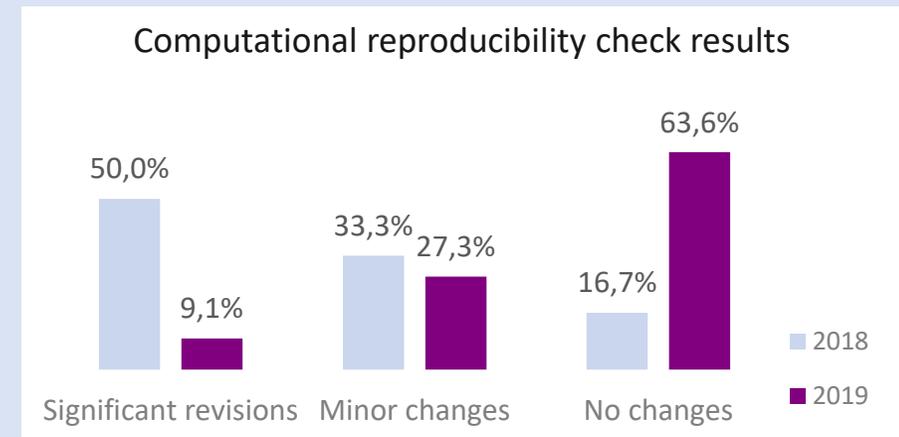
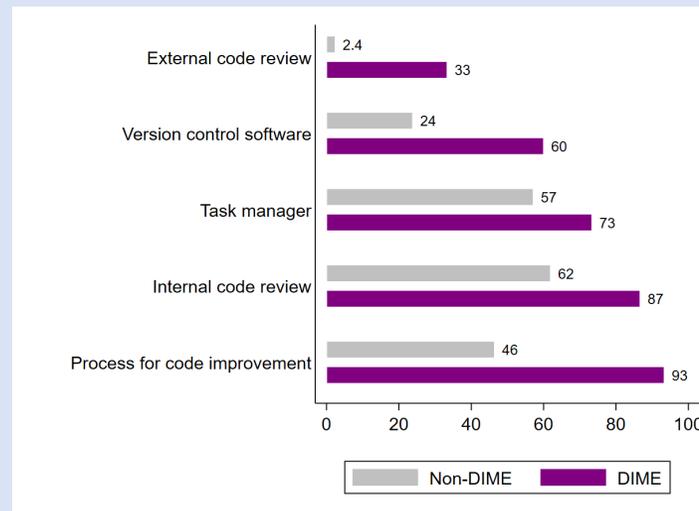
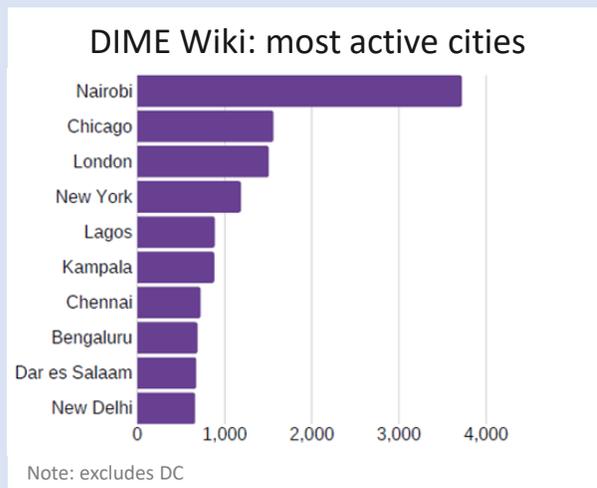
Revisione del sondaggio/Libreria dei questionari/

Ricerca riproducibile

"Bootcamp" e revisione trimestrale del codice peer

Controlli di riproducibilità computazionale al 100%

Supporto per la pubblicazione dei dati



Beni pubblici per la ricerca



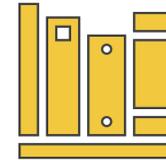
[Development Research in Practice](#)

The DIME Analytics Data Handbook



[DIME Wiki](#)

One-stop shop for impact evaluation resources



[Stata visual library](#)

Reproducible Stata code for graphs and maps



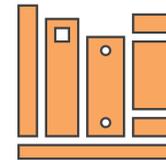
[iefieldkit](#)

Tools for primary data collection and processing



[ietoolkit](#)

Tools for data management and analysis



[R econ visual library](#)

Reproducible R code for impact evaluation data visualization

Valutazione d'impatto

- La valutazione dell'impatto della politica pubblica è un buon principio per l'organizzazione dei dataset di alta qualità
- Ma i dati amministrativi sono disegnati tipicamente per la produzione di statistiche ufficiali e il seguimiento dei processi amministrativi: *questo riduce il loro uso per l'analisi della politica*
- @DIME combiniamo un approccio di sviluppo delle capacità governative per accesso ai dati amministrativi con l'integrazione con altri dati sia nuovi che esistenti nella creazione di dataset di alta qualità e pronti per l'analisi

- {DATI} Innovare nella misurazione e costruire set di dati per comprendere il problema che stiamo cercando di risolvere
- {AI} Analizzare i dati e trovare opportunità per direzionare l'azione e prioritizzare investimenti e politiche
- {IE} Inferenza causale: sperimentazione via studi randomizzati controllati e valutazione di investimenti e politiche per aumentare l'efficacia, documentare il successo e giustificare lo scale up

MIGLIORARE

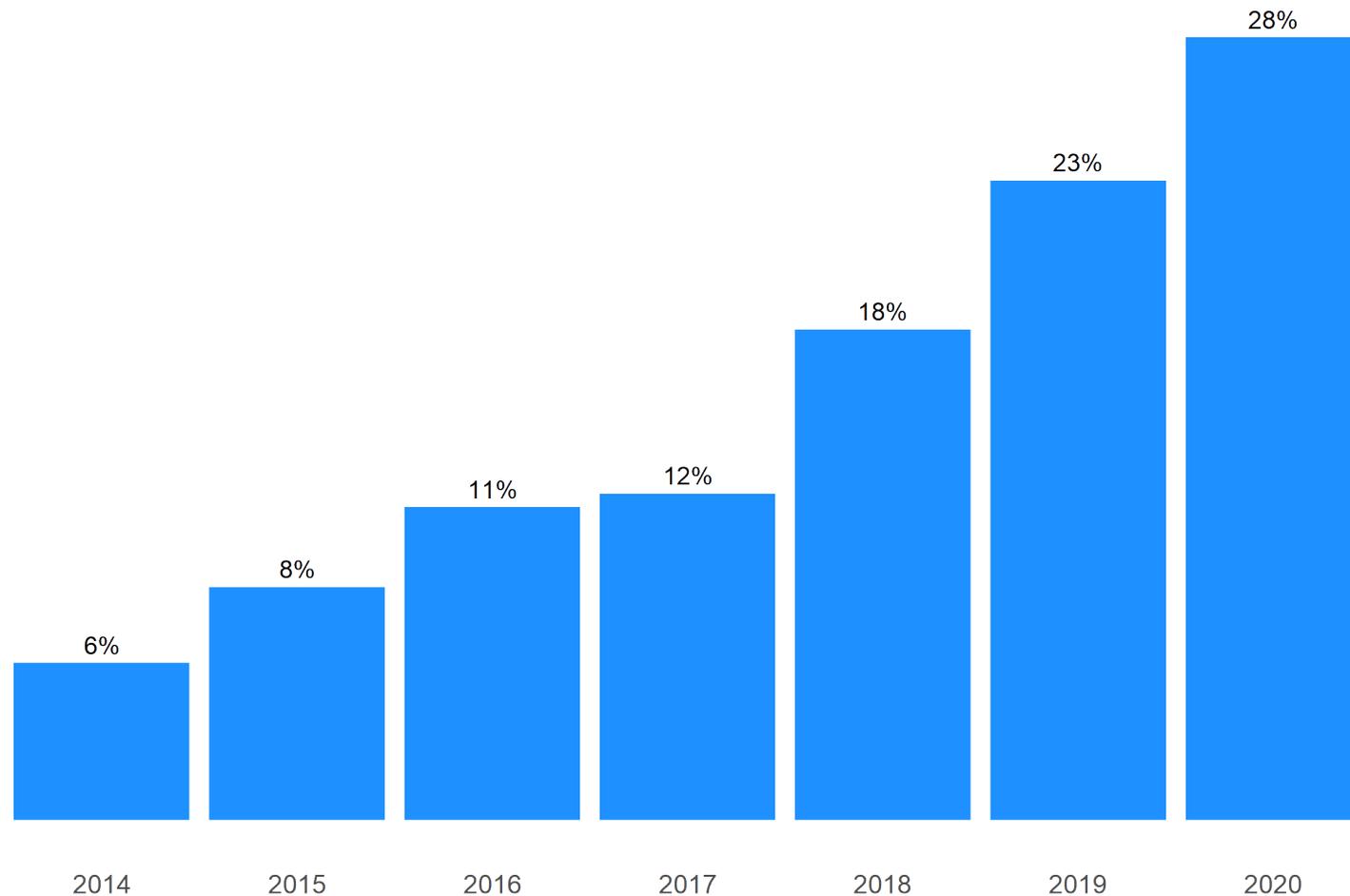
- PIANIFICAZIONE
- SEGUIMENTO
- VALUTAZIONE
- RISULTATO

Per il settore pubblico, l'uso dei dati amministrativi può

- Ridurre drammaticamente il costo della ricerca
 - Aumentare frequenza, copertura, quantità dei dati
 - Ridurre attrito
 - Aumentare la frequenza e le braccia degli studi randomizzati controllati
- Aumentare la frequenza e disponibilità di evidenza per le decisioni di politica economica

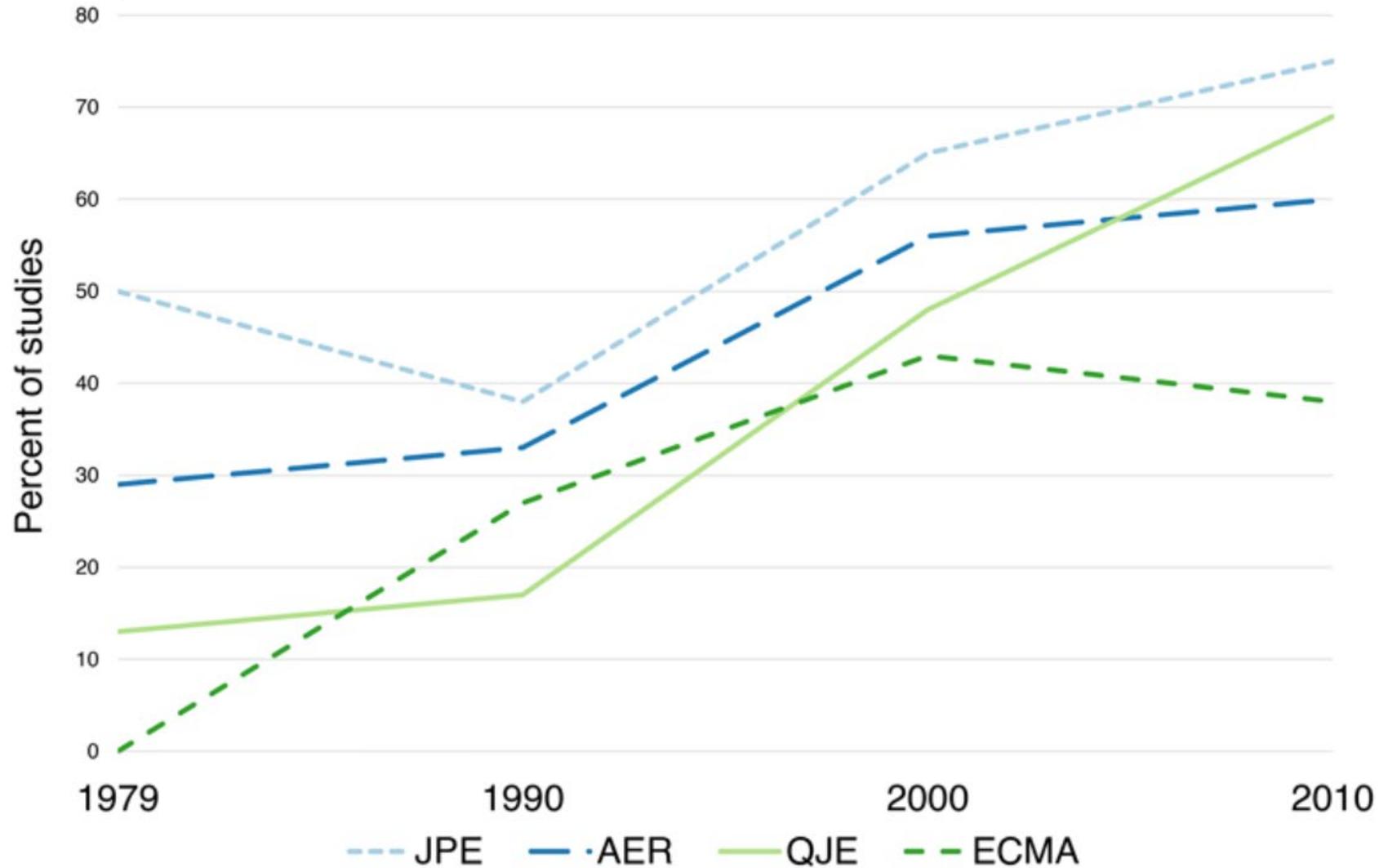
L'uso dei dati amministrativi cresce storicamente (DIME)

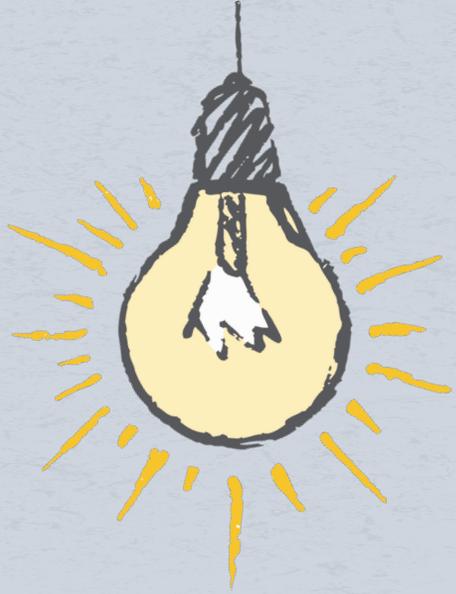
DIME Impact Evaluations Using Administrative Data
Percent of active projects that used administrative data



Percentuale di studi condotti in paesi ad alto reddito che utilizzano dati amministrativi, tra gli studi pubblicati nelle quattro principali riviste statunitensi in Economia (Journal of Political Economy, American Economic Review, Quarterly Journal of Economics, Econometrica).

Fonte: Chetty (2012).





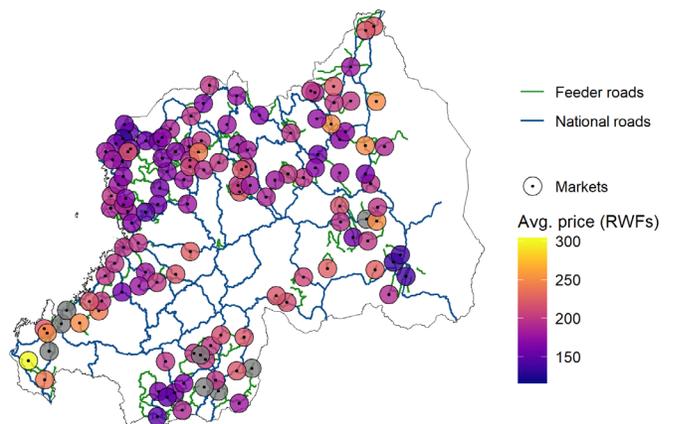
CASO 1. Sviluppo sistemi di dati nazionali per la ricerca economica : programma Rwanda

Florence Kondylis, Maria Jones, John Loeser (DIME)
and Jeremy Magruder (U Berkeley)

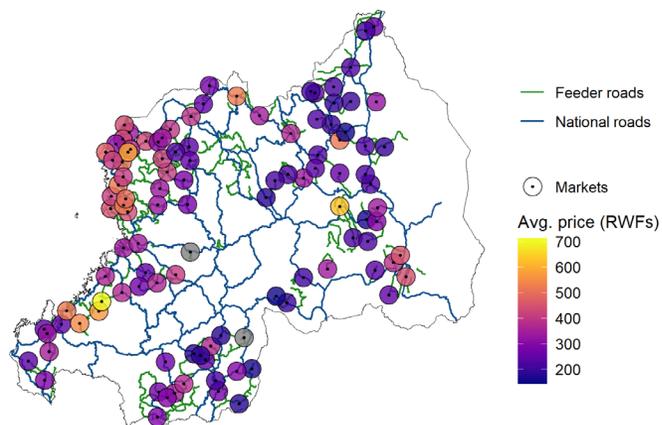
Utilizzo di dati amministrativi esistenti combinati con e.g., dati sui prezzi di mercato ad alta frequenza aumenta la portata dell'analisi



Raccogliamo in remoto informazioni mensili su 68 prodotti in 150 mercati

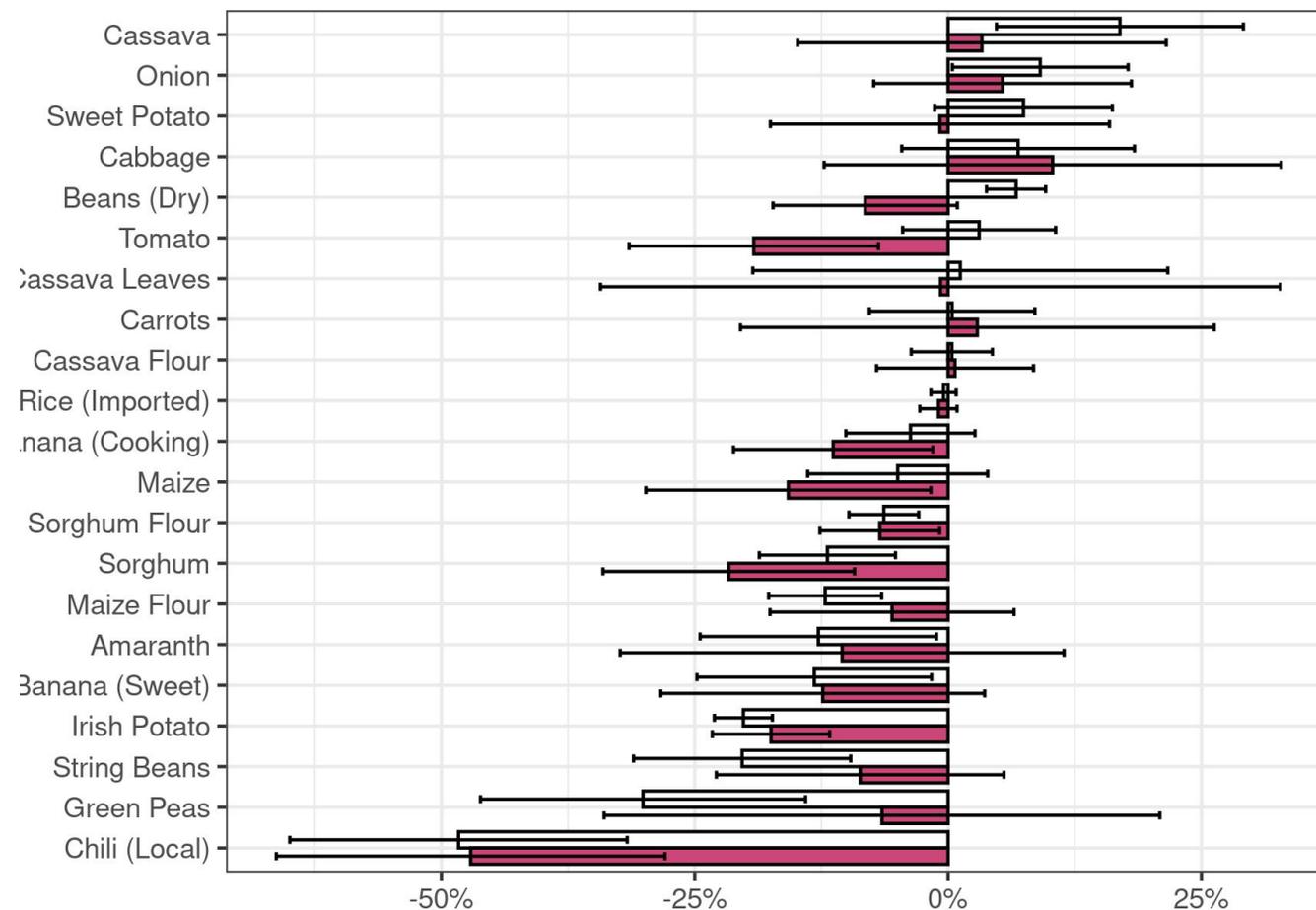


Potato prices



Tomato prices

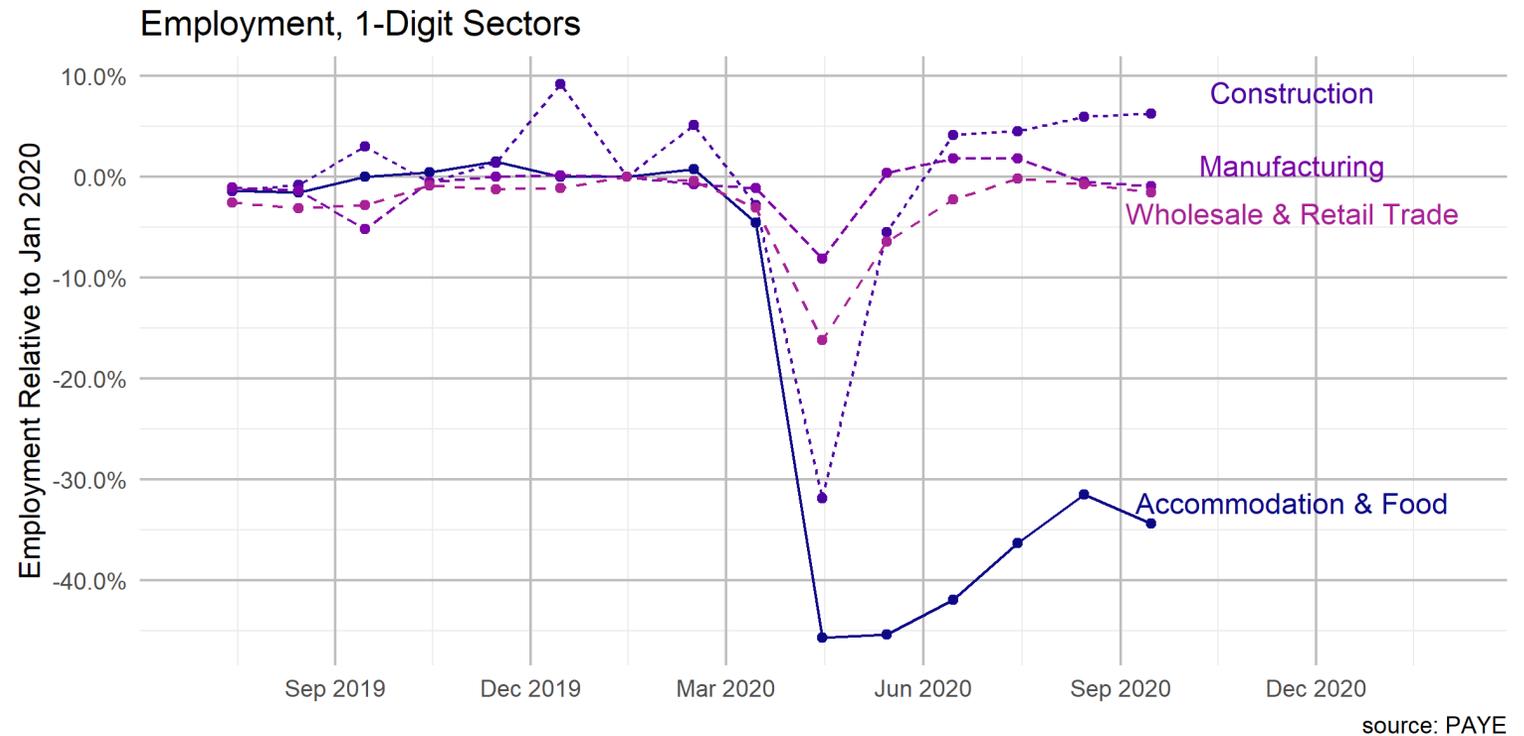
□ April ■ May



Percent change in seasonally adjusted prices between April/May and Feb 2020 (relative to same months in 2018 and 2019)

Ruanda: dati fiscali di fatturazione elettronica per approfondimenti di politica economica in tempo reale

- In collaborazione con la Rwanda Revenue Authority, costruito un dashboard per il monitoraggio mensile delle entrate fiscali settoriali durante COVID-19
- Mentre l'economia generale si è ripresa rapidamente alla fine del 2020, il dashboard ha rivelato che i settori in cui erano necessarie interazioni faccia a faccia con i consumatori hanno continuato a registrare una ripresa prolungata
- La politica fiscale durante il COVID-19 non può essere indipendente dal settore



K Bryne, S Karpe, F Kondylis, M Lang, J Loeser. *Sectoral heterogeneity in the COVID-19 recovery: Evidence from Rwanda*. Featured in R Arezki, S Djankov, and U Panizza "Shaping Africa's Post-Covid Recovery" (CEPR/VoxEU, Feb 2021).

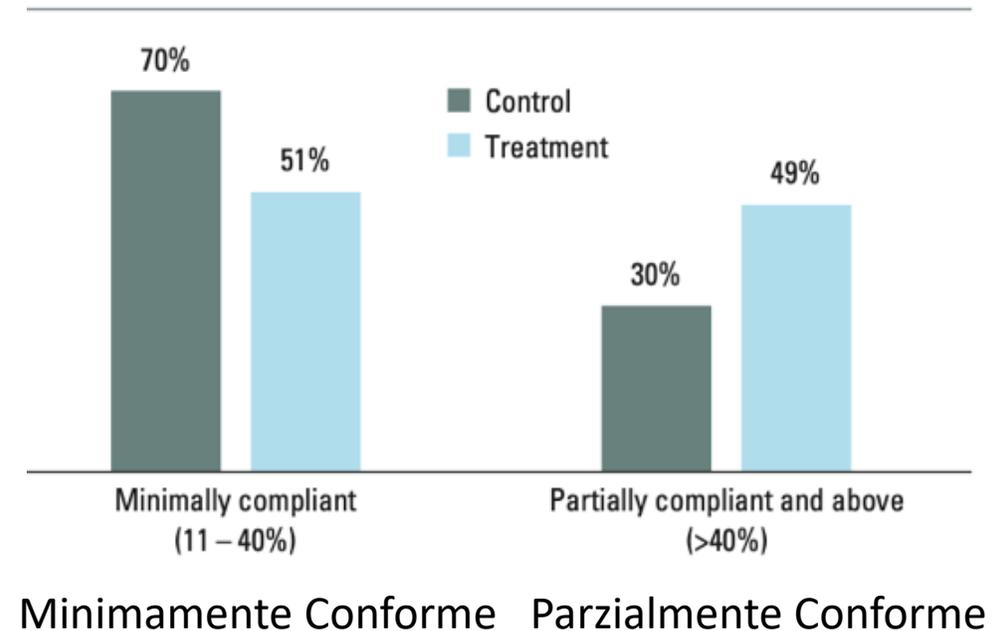


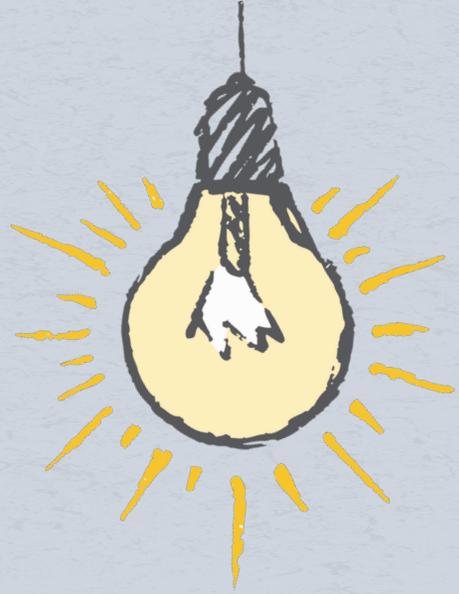
Caso 2. Pilotare nuovi sistemi di dati amministrativi per stabilire sistemi di «accountability»: sicurezza dei pazienti in Kenya

Guadalupe Bedoya (DIME)
Jishnu Das (Georgetown U.)

Valutazione dell'impatto sulla sicurezza dei pazienti in Kenya

- Non erano disponibili dati sistematici sulla conformità agli standard di sicurezza dei pazienti nelle strutture sanitarie
- I ricercatori DIME hanno pilotato un sistema di ispezione sanitaria ad alto rischio standardizzato in collaborazione con il Ministero della Salute
- Lo studio ha dimostrato che il regime delle ispezioni con regole chiare, monitoraggio e applicazione di premi e punizioni è efficace nel migliorare la sicurezza dei pazienti nelle strutture sanitarie pubbliche e private
- Prova del concetto del valore di sostenere i governi nello sviluppo di sistemi di dati per valutare le riforme a livello di settore e stabilire sistemi di *accountability*





Caso 3. Digitalizzazione dei dati amministrativi e sviluppo sistemi di Raccolta digitale dei dati al punto di origine: sicurezza stradale in Kenya

Guadalupe Bedoya, Arianna Legovini, Rob Marty, Sveta Milusheva (World Bank)
Sarah Williams (MIT)

<https://journals.plos.org/plosone/article/comments?id=10.1371/journal.pone.0244317>

Nairobi: Mapping police crash data

- Ex-post digitizzazione: Scansionato tutti i rapporti di incidenti da tutti i 15 distretti di polizia di Nairobi (oltre 10.000 rapporti)
- Rapporti digitalizzati, comprese le caratteristiche dell'incidente e la geoposizione dell'incidente
- Circa 9.000 incidenti unici geo-identificati dal 2011 al 2018

Time

Date

Injury severity

Location

Direction of travel

Vehicle type

Cause code of crash used by police

Hospital taken to

Number of vehicles and Injuries

REF. TC/1/14/INDO/42./2017 27/6/2017 PD

SUBJECT SELF INVOLVED SERIOUS INJURY ROAD TRAFFIC ACCIDENT PD THIS OCCURED TODAY

27/6/2017 ABOUT 01.00 AM AT LUSAKA ROUNDABOUT INVOLVING M/V REG. NO. [REDACTED] V/W

PASSAT DRIVEN BY [REDACTED] PD IT HAPPENED THAT

THE SAID M/V WAS BEING DRIVEN FROM J.K.I.A. DIRECTION HEADING TOWARDS CITY CENTRE AND ON

REACHING AT YTHE LOCATION OF THE ACCIDENT , THE DRIVER LOST CONTROL OF THE M/V , AND HIT

THE PAVEMENT AND LANDED INTO ROUNDABOUT PD DUE TO IMPACT THE M/V GOT FIRE PD AS A

RESULT THE DRIVER SUSTAINED FRACTURES ON BOTH LEGS AND WAS TAKEN TO K.N.H. WHERE HE WAS

LEFT UNDERGOING TREATMENT PD SCENE VISITED AND M/V SHELL TAKEN TO THE STATION PD CAUSE

CODE - 26 USED PD CASE P.U.I. , D.T.O. INDO/AREA DEALING PD

BT

DRAFTER – PC KANDIE

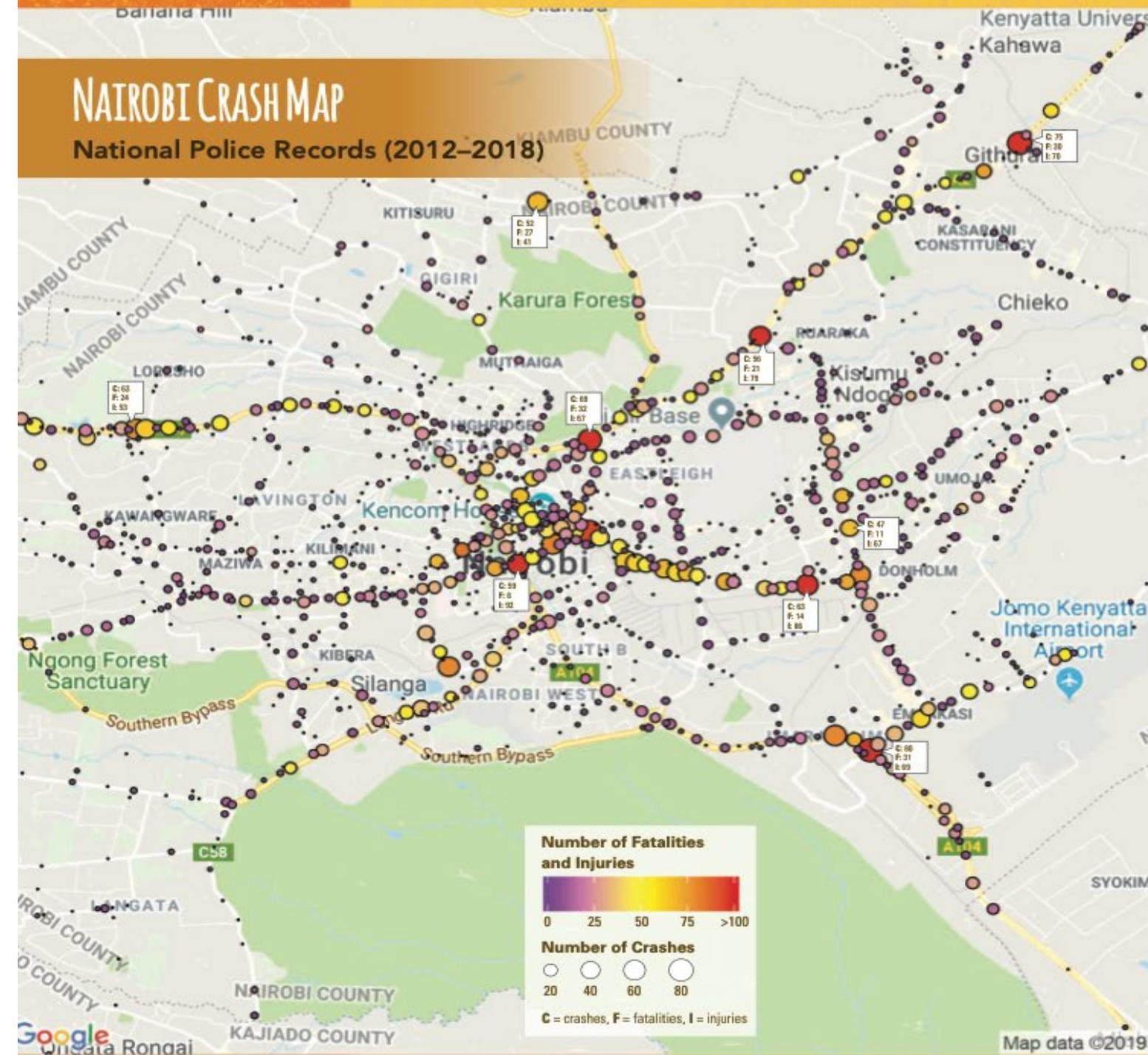
FOR D.T.E.O. INDO/AREA



[Link to Dashboard](#)

NAIROBI CRASH MAP

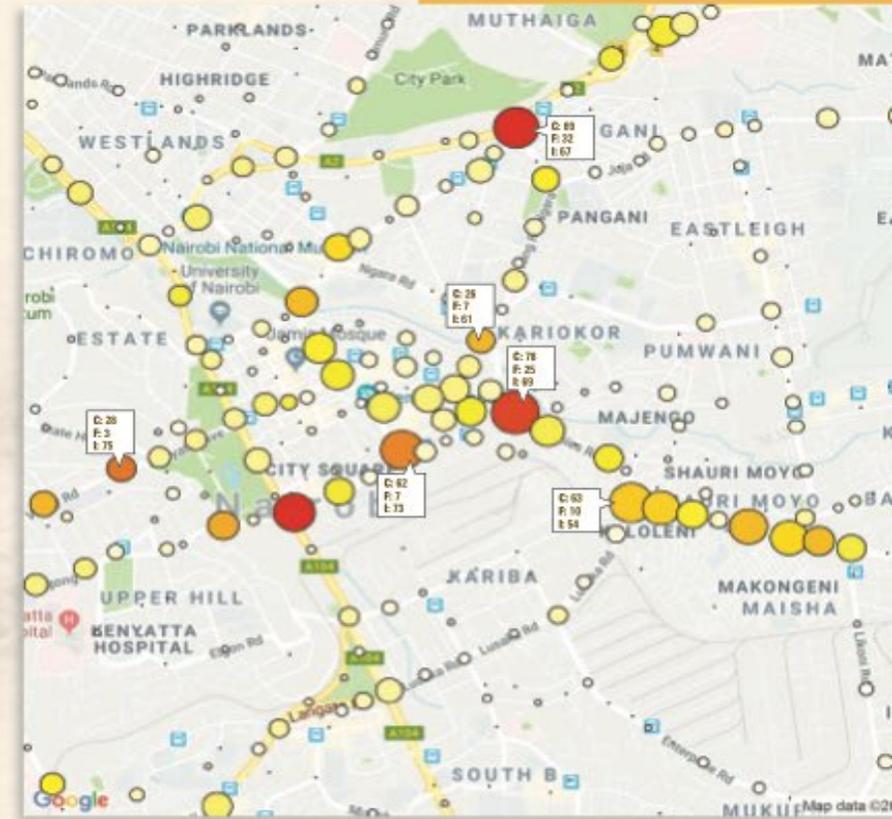
National Police Records (2012–2018)



71% of road traffic deaths are **pedestrians**



CENTRAL BUSINESS DISTRICT



200 locations (blackspots) represent **52%** of crashes and **55%** of deaths

Blackspots are locations where a high number of crashes, deaths or injuries have occurred in a delimited space and time.

INNOVAZIONI AI PER GEOLOCALIZZARE POST SUI SOCIAL MEDIA CHE RIPORTANO INCIDENTI STRADALI (CON INTEGRAZIONE DATI TWITTER, WEATHER, GOOGLE MAPS, UBER)

Ma3Route @Ma3Route · 1h
 12:37 bad **accident** on **waiyaki way** **next to kianda** heading towards **abc place**

Classify Reports

```
{bad} {accident} {on}
{waiyaki} {way} {next} ...
{bad accident} {accident on} {on waiyaki}
{waiyaki way} {way next} ...
{bad accident on}
{accident on waiyaki}
{on waiyaki way} ...
```

Support Vector Machine → Crash Report ✓

Extract landmarks and roads with exact match to gazetteer, misspells and abbreviations

Landmarks: kianda, agent kianda, kianda college, kianda electrical, kianda school parking, kianda secondary school, abc place, chase bank abc place, dr mattress abc place, abc place parking

Roads: waiyaki way

Map:

Choose landmark based on prepositions

(next to - tier 1)
(towards - tier 6)

kianda, agent kianda, kianda college, kianda electrical, kianda school parking, kianda secondary school

waiyaki way

Map:

Restrict to landmarks within 500m of road

kianda school parking, kianda secondary school

waiyaki way

Map:

Use centroid of landmarks

centroid of: {kianda school parking, kianda secondary school}

waiyaki way

Map:

Extract all crash reports within 500 meters reported within 4 hours of each other

12:37 bad accident on waiyaki way next to kianda heading towards abc place

12:50 pedestrian hit just before kebete lab

12:23 wreck near sky park - traffic at standstill

14:45 accident near kianda school cleared, traffic flowing

13:15 traffic on waiyaki way wreck near afrati guest house

Map:

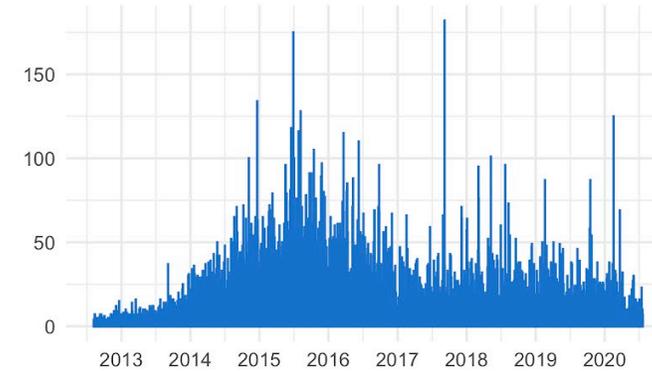
Cluster Crash Reports into Crash

Take centroid of locations and earliest crash report

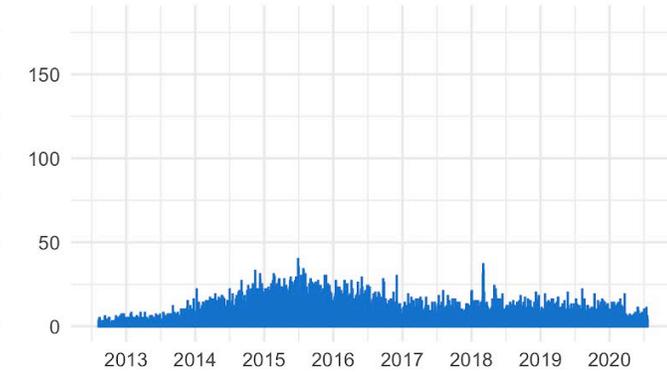
Crash Location:

Crash Time: 12:23

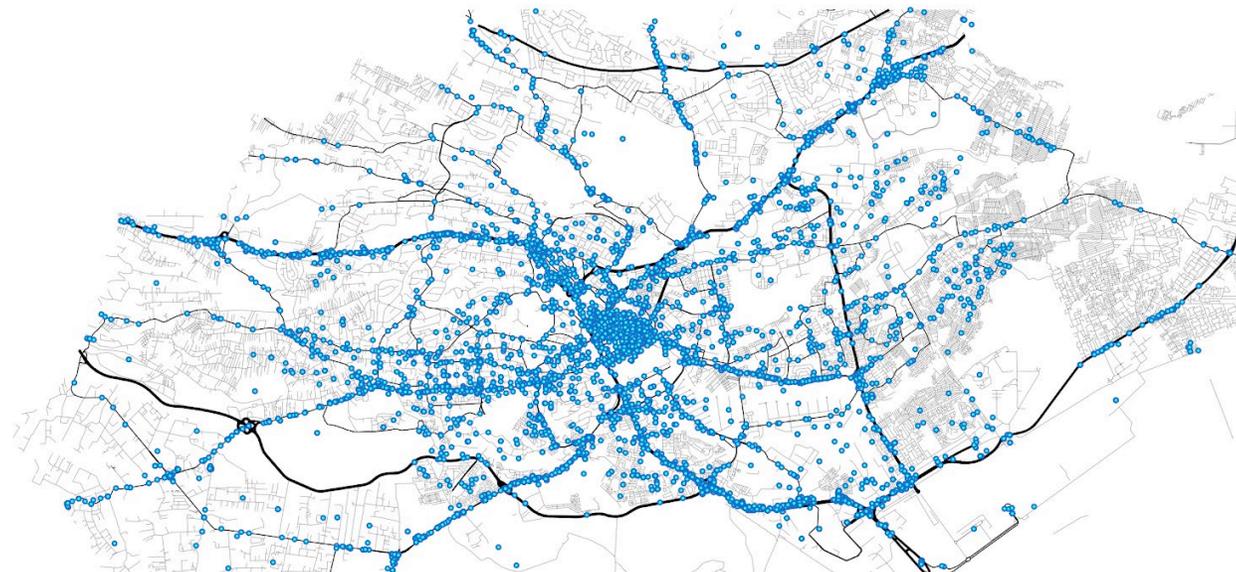
A Daily Number of Tweets that Report a Crash



B Daily Number of Geolocated Crashes

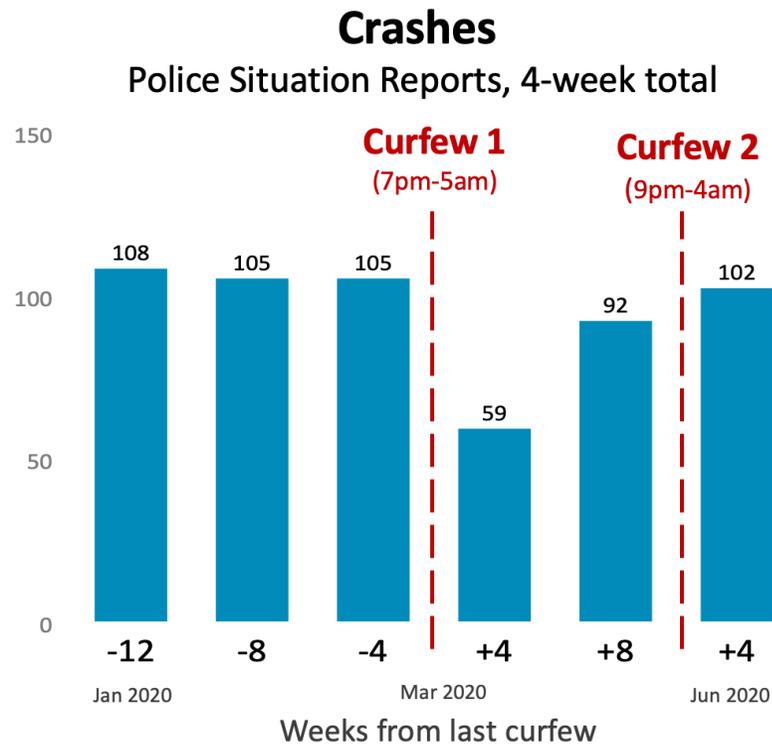


C Crashes, August 2012 - July 2020



- Raccolta digitale dei dati al punto di origine: sviluppo sistema E-crash
- Utilizzo di dati ad alta frequenza per dare priorità agli investimenti nella sicurezza stradale

- Creare sistemi di dati pilota sugli incidenti stradali
- Produci analisi sui punti critici di crash e sui rischi associati
- Generare conoscenza sull'impatto degli interventi di sicurezza stradale





Caso 4. Sfruttare gli investimenti nei sistemi di dati amministrativi: DE JURE

Daniel Chen et al. (World Bank)

DE JURE

Le recenti innovazioni hanno aperto nuove opportunità per la giustizia:

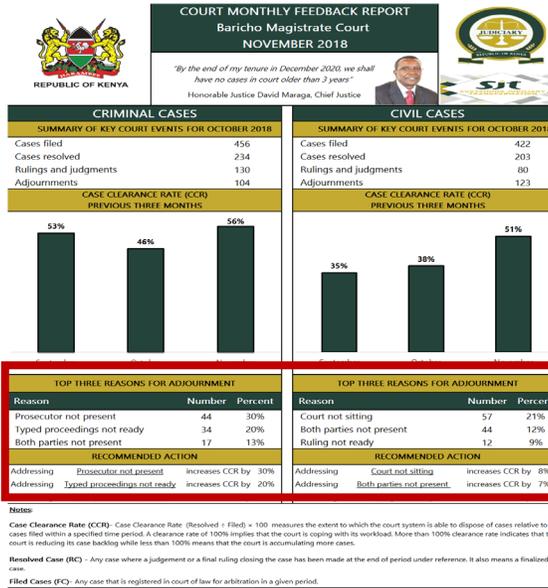
- Dataset su larga scala sempre più digitalizzato
- Applicazioni ML per produrre dati interpretabili da testo non strutturato
- Modelli predittivi del processo decisionale per comprendere meglio i pregiudizi e affrontarli con interfacce digitali



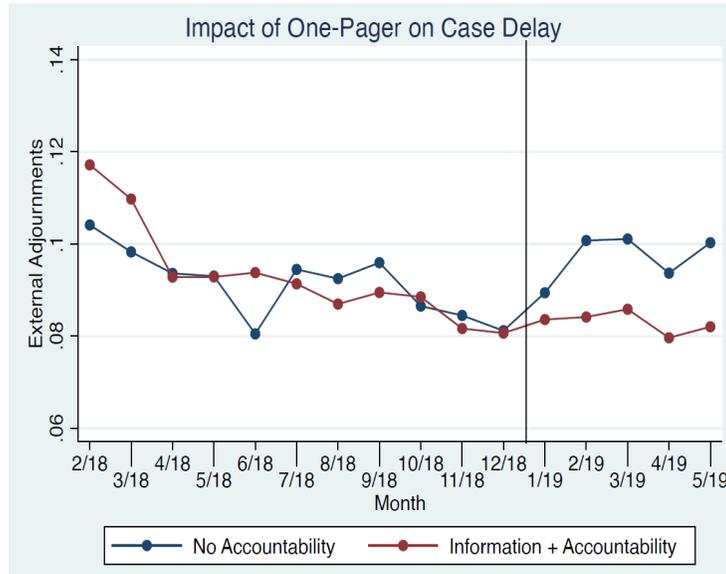
Innovazioni e-justice

AI:
Algoritmi per identificare le maggiori fonti di ritardo giudiziario

Information and accountability



Increased speed of justice

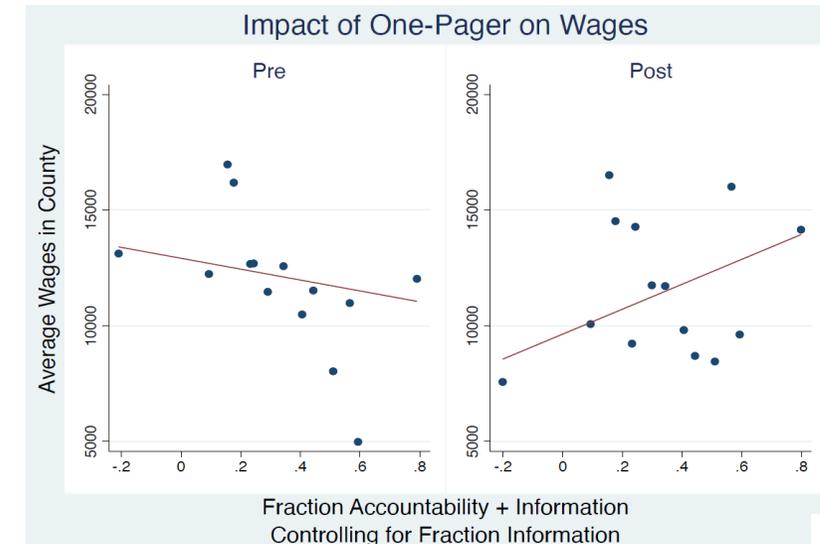


Identifying key reasons for adjournments

TOP THREE REASONS FOR ADJOURNMENT		
Reason	Number	Percent
Prosecutor not present	44	30%
Typed proceedings not ready	34	20%
Both parties not present	17	13%

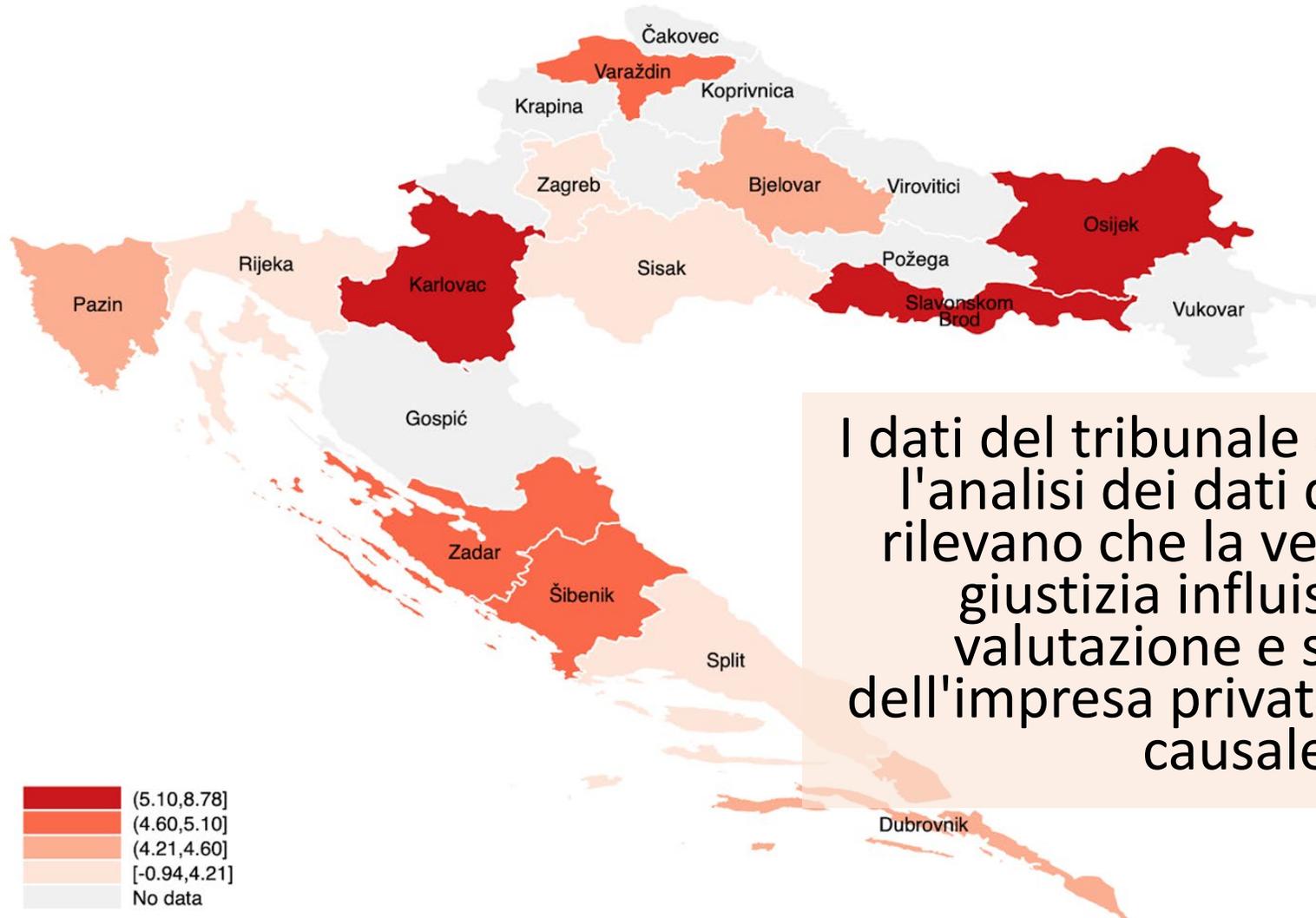
RECOMMENDED ACTION		
Addressing <u>Prosecutor not present</u>	increases CCR by	30%
Addressing <u>Typed proceedings not ready</u>	increases CCR by	20%
Addressing <u>Both parties not present</u>	increases CCR by	13%

Increased wages & economic outcomes



Elasticity of Firm Revenue to Judicial Speed in Commercial Courts

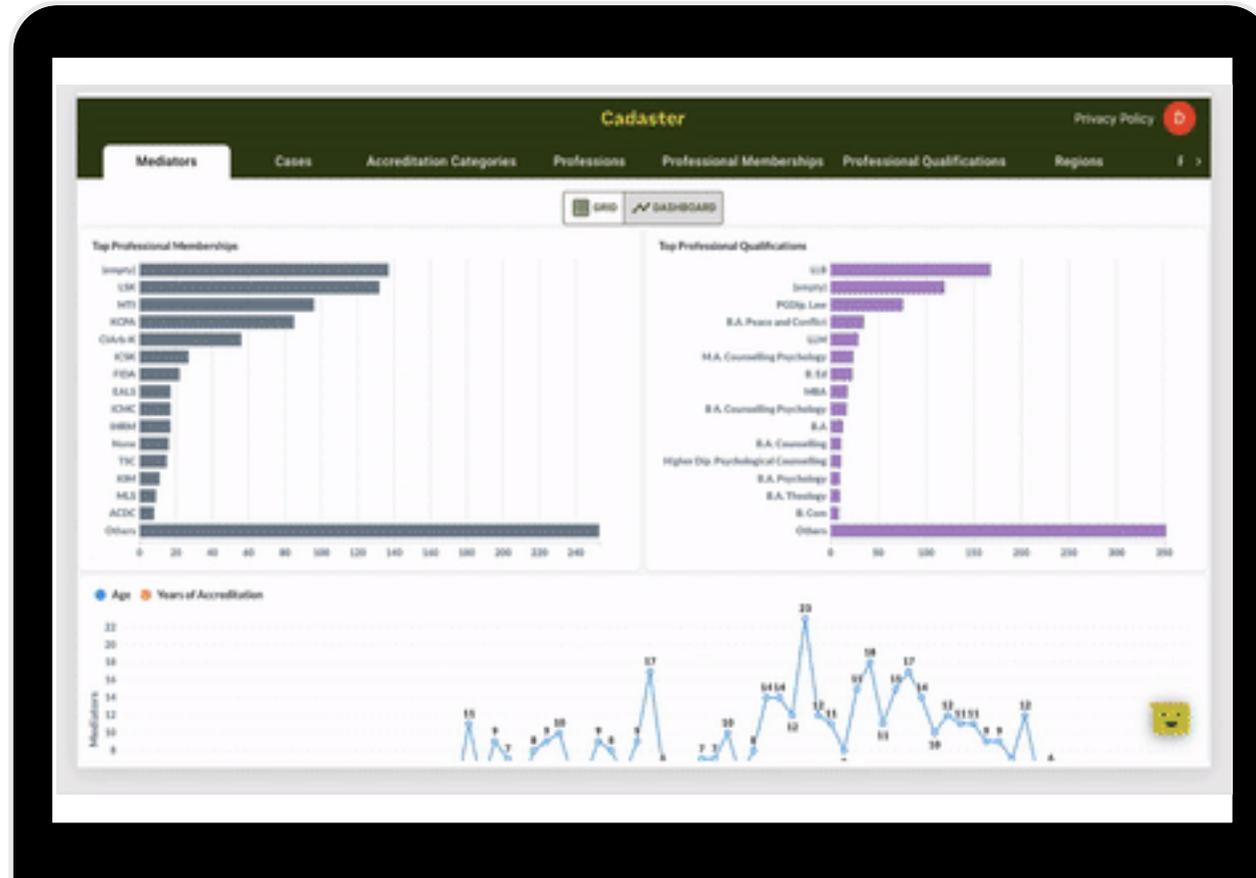
Years: 2010-2016



I dati del tribunale elettronico + l'analisi dei dati di mercato rilevano che la velocità della giustizia influisce sulla valutazione e sui ricavi dell'impresa privata in maniera causale

Piattaforma di Mediazione Legale

- Piattaforma ICT per mediatore e gestione dei casi
- Assegnazione intelligente dei casi ai mediatori
- Gestione delle prestazioni e avvisi in tempo reale
- Strumenti open source senza codice per l'inserimento dei dati e il supporto alle decisioni



DIGITIZE

MANAGE

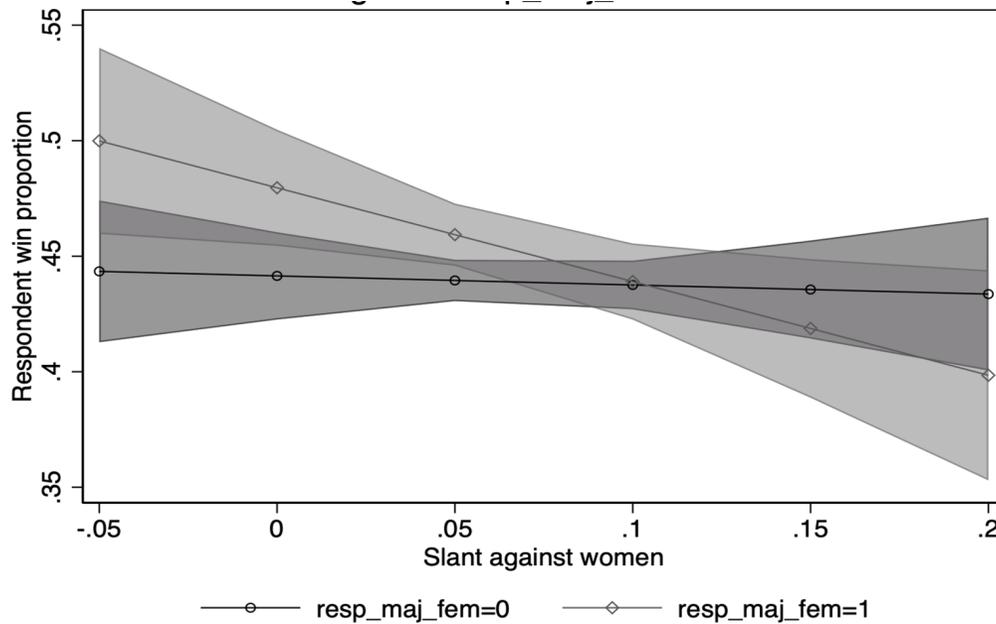
TRACK

OPTIMIZE

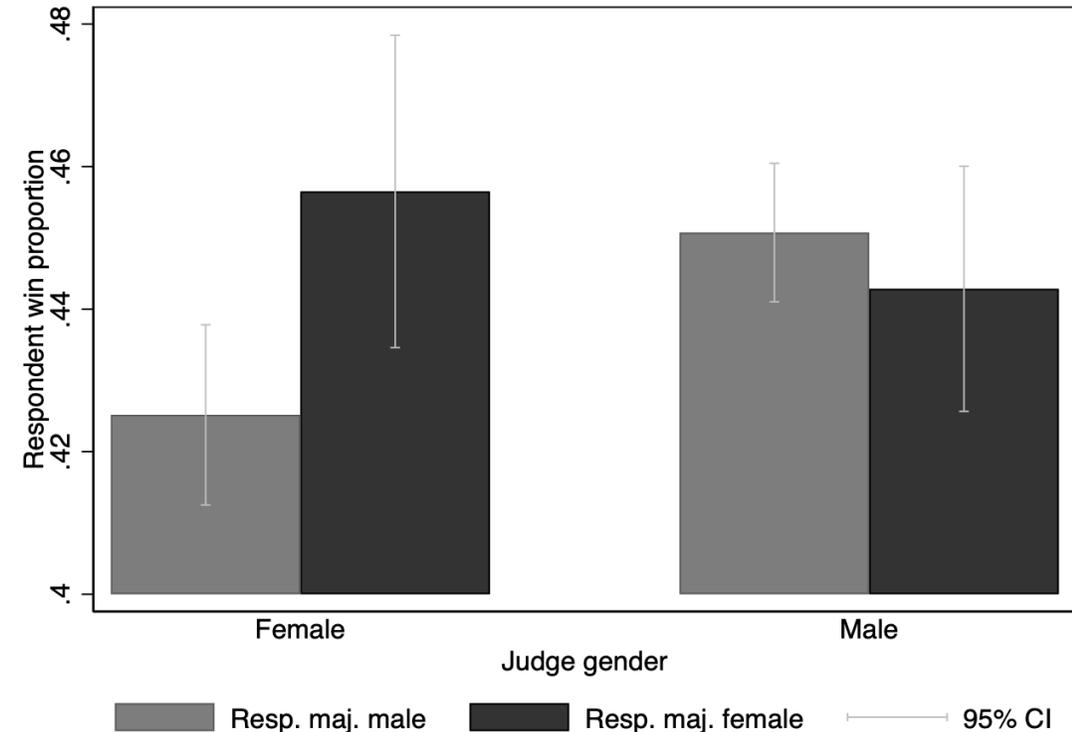
AI per misurare e affrontare i pregiudizi di genere, etnici e religiosi nelle decisioni giudiziarie

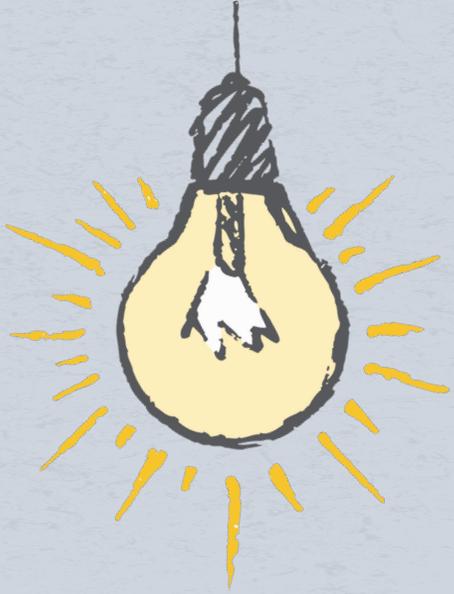
- Contributi analitici: sfruttare i sistemi di dati e l'apprendimento automatico per migliorare la qualità della giustizia.
- Identificazione dei pregiudizi: gli esiti giudiziari dipendono dal genere/etnia?
- Affrontare attraverso incentivi: come ridurre i pregiudizi giudiziari?

D: I pregiudizi di genere sono correlati con le decisioni contro le parti in causa? Sì, i giudici con più pregiudizi hanno maggiori probabilità di pronunciarsi contro le imputate



D: Il genere del giudice influisce sull'imparzialità della giustizia in Kenya? Sì, i giudici tendono ad essere più indulgenti con gli intervistati/accusati del proprio genere



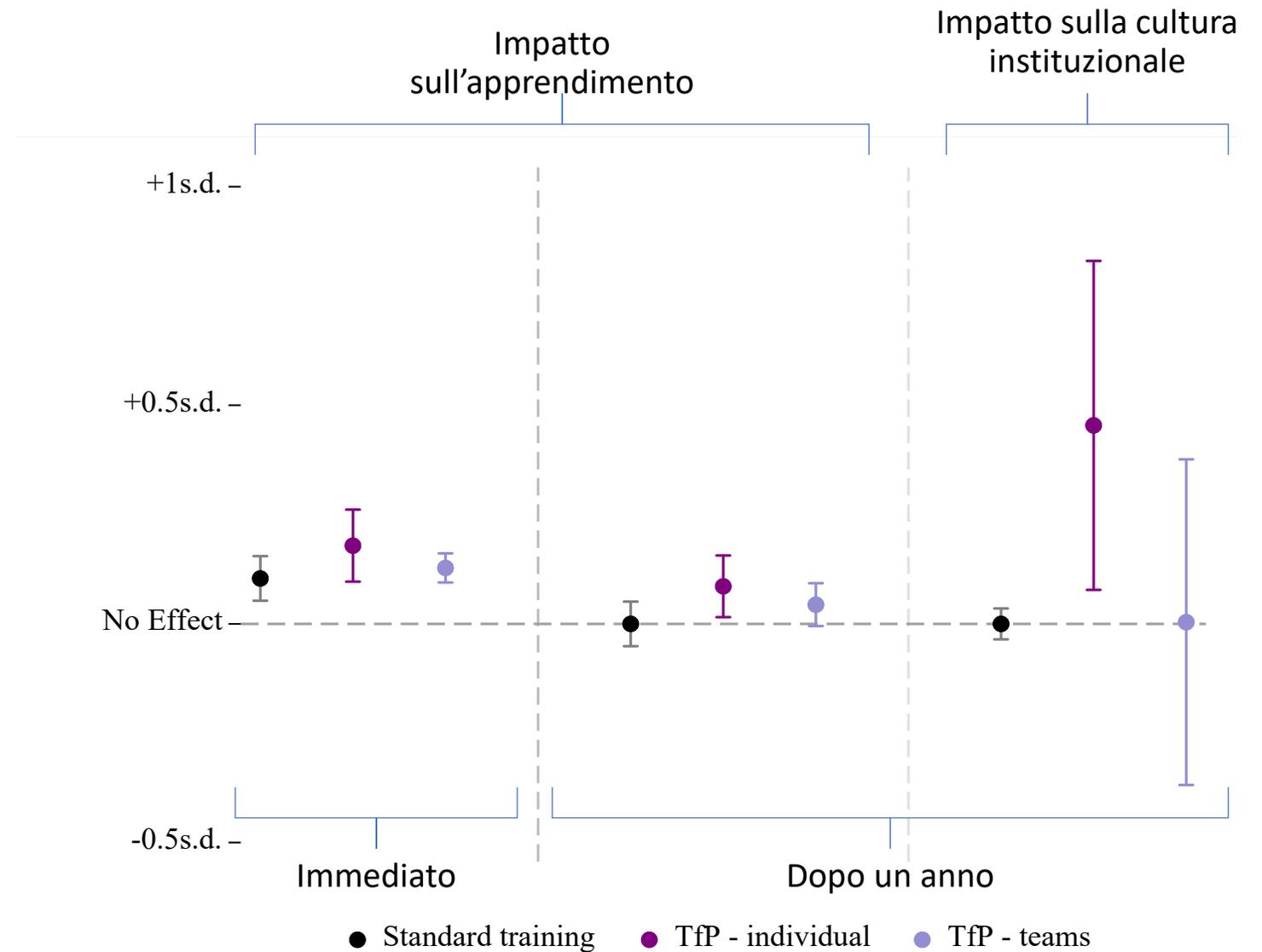


Caso 5. Sviluppare inchieste del servizio civile: Bureaucracy Lab

Dan Rogger et al. (DIME)

Costruire la capacità del servizio pubblico: il programma di formazione per la produttività del Ghana

- RCT aggiungendo il modulo "Formazione per la Produttività" alla formazione standard all'interno del servizio civile aumentata persistenza dell'apprendimento mediato da cambiamenti culturali



Learn more (links)

- [DIME brochure](#)
- [DIME Annual Report](#)
- [DIME Website](#)
- [Development Impact Blog |](#)
- [DIME Wiki](#)



SCIENCE FOR IMPACT
 BETTER EVIDENCE FOR BETTER DECISIONS
 ————— THE DIME EXPERIENCE —————

Impact Evaluation Helps Deliver Development Projects

*Arianna Legovini
 Vincenzo Di Maro
 Caio Piza*

Development Impact
 Laws, views, methods, and insights from the world of evaluation

Tags | Contact

Regular 3

on spending
 ident learning
 tes

Work: Shhh!
 s, and
 iment: Do
 'esponse
 thods
 'ence

How to Implicit Association Test?

SUBMITTED BY FLORENCE KONDYLIS ON WED, 04/10/2019
 CD-AUTHORS: LUIZA ANDRADE, ARIANNA LEGOVINI, KATE VYBORNY, ASTRID ZWAGER

How to do Implicit Association Test?
 Implicit Association Tests (IATs) are being increasingly used in applied m papers. While IATs can be found off-the-shelf, designing your own IAT r you to get at respondents' implicit attitudes towards something more We added a custom IAT to a survey of commuters in Rio de Janeiro, a go over the practical steps involved. For our project, we wanted to r and female commuters' implicit attitudes towards women riding r' the co-ed car relative to women riding the women's-only car. Th quantify the stigma women may face for not using gender-se

Read more =>

discrimination | Field Survey | IAT | Tr

...valuation (DIME)

THE "PROGRAM YOUR FU...
Bridging the gap
 April 11, 2019 — DIME is testin...
 Brazil's most vulnerable youth...
 blockchain programming acad...
 happen! [Read More >](#)

OUR WORK

Agriculture
 Transforming rural space through...
 technology, measurement and...
 supply chains to secure brighter...
 and more commercial futures.

Energy and Environment
 Are we doing enough on climate?...
 No. From satellite data to farmer...
 incentives, speeding up learning on...
 what's possible.

Entertainment Education
 What public campaigns cannot...
 professional entertainers can...
 amazing power of the narrativ...
 changing both norms and...
 behaviors.