This document facilitates replication of the data formation and analysis in *Under the Cover of Darkness: How Ambient Light Influences Criminal Activity*. We describe below how to construct our final dataset and where to find relevant original data. Users interested in replicating results without building data from scratch can use the dataset titled "maindata" available on the RESTAT website.

1 Original data

1.1 NIBRS

Data source: http://www.icpsr.umich.edu/icpsrweb/NACJD/NIBRS/

Data files: nibrs2005_1.dta, nibrs2005_2.dta, nibrs2005_3.dta, nibrs2005_4.dta, nibrs2005_5.dta, nibrs2005_6.dta, nibrs2005_7.dta, nibrs2006_1.dta, nibrs2006_2.dta, nibrs2006_3.dta, nibrs2006_4.dta, nibrs2006_5.dta, nibrs2006_6.dta, nibrs2006_7.dta, nibrs2007_1.dta, nibrs2007_2.dta, nibrs2007_3.dta, nibrs2007_4.dta, nibrs2007_5.dta, nibrs2007_6.dta, nibrs2007_7.dta, nibrs2008_1.dta, nibrs2008_2.dta, nibrs2008_3.dta, nibrs2008_4.dta, nibrs2008_5.dta, nibrs2008_6.dta, nibrs2008_7.dta

Additional information:

Data on crime are from the National Incident-Based Reporting System (NIBRS) for years 2005-2008. To create a balanced panel, we restrict our attention to the 558 jurisdictions that consistently report across all four years.

A number of regions report an abnormally large number of crimes at midnight of each day, which we take as an indicator of "bunching" in reporting. Given our interest in specific recorded hours, we drop all jurisdictions with a modal reporting hour of midnight (prior versions of this paper showed including these jurisdictions did not change our daily results where hourly bunching is less concerning). We also drop jurisdictions without at least one reported incidence of robbery, rape, and assault per year and jurisdictions from the state of Arizona, which does not follow DST (and has only two reporting jurisdictions).

As additional data cleaning, we omit data from a Henrico County, a jurisdiction in Virginia, where reported crime data had a number of errors (jurisdiction identifier number VA0430100). We also omit data where the reported hour is missing (coded as "99").

NIBRS only report hours in which at least one occurrence of monitored crimes. For example, if no reported crime occurred at 6 pm on a given day in a given jurisdiction, a data entry for that jurisdiction/hour would not exist. To address this issue, we expand the data to cover all hours for every jurisdiction and and enter values of zero for reporting in crimes with no hours.

1.2 Matching ORIs with states and counties

Data source: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4634

Data file: crosswalk_stata_version.dta

Additional information:

This is a crosswalk file, .

1.3 Timezones

Data source: http://www.nws.noaa.gov/geodata/catalog/wsom/html/cntyzone.htm

Data file: fips_timezones.csv

Additional information:

This crosswalk matches jurisdictions to relevant time zones.

1.4 Sunrise and sunset times

Data source: http://www.esrl.noaa.gov/gmd/grad/solcalc/

Data file: latlong_sunrise.dta

Additional information:

We collected "apparent sunrise" and "apparent sunset" times for each relevant date and latitude-longitude pair from the National Oceanic & Atmospheric Association website (http://www.esrl.noaa.gov/gmd/grad/solcalc/).

Though sunset information is down to the minute, crime data are only available by hour, so we simplify the relevant sunset time using the following strategy:

- Find the hour and minute of sunset for the day before DST, using the latitude and longitude data.
- Generate a "time since prior sunset" variable, equal to the hour of the reported crime minus the hour and minute of sunset found above.
- Round the "time since prior sunset" variable to the nearest hour.

As an example, take a jurisdiction in a region with a sunset time of 6:23 PM on the day before DST begins. Crimes reported as occurring between the hours of 6 and 7 PM will be assigned a "time since prior sunset" of 6 PM − 6:23 PM = −23 minutes, which will round to 0 hours. A crime recorded in the 8-9 pm hour will be given a value of 2 hours. As sunset time varies across the United States, crimes in the 8 PM hour can have different "time since prior sunset" depending on their geographic location.

1.5 Weather data

Data source: Wolfram Schlenker, http://www.wolfram-schlenker.com/dailyData.html

Data file: Data file: county_weather_2005_2008.dta

Additional information:

Wolfram Schlenker generously provided weather data for this project — interested parties should contact Prof. Schlenker for data availability. We provide a Stata do-file for analysis both with and without these data. Should you choose to use these or similar weather data, we include average temperature and precipitation.

2 Creation of final dataset

Do-file: doleac_sanders_replicate.do

Input data files: fips_timezones.csv, crosswalk_stata_version.dta, latlong_sunrise.dta, nibrs2005_1.dta, nibrs2005_2.dta, nibrs2005_3.dta, nibrs2005_4.dta, nibrs2005_5.dta, nibrs2005_6.dta, nibrs2005_7.dta, nibrs2006_1.dta, nibrs2006_2.dta, nibrs2006_3.dta, nibrs2006_4.dta, nibrs2006_5.dta, nibrs2006_6.dta, nibrs2006_7.dta, nibrs2007_1.dta, nibrs2007_2.dta, nibrs2007_3.dta, nibrs2007_4.dta, nibrs2007_5.dta, nibrs2007_6.dta, nibrs2007_7.dta, nibrs2008_1.dta, nibrs2008_2.dta, nibrs2008_3.dta, nibrs2008_4.dta, nibrs2008_5.dta, nibrs2008_6.dta, nibrs2008_7.dta, county_weather_2005_2008.dta

Output data file: maindata.dta

3 Analysis

Do file: doleac_sanders_replicate.do

Additional information:

The provided do-file provides the code needed to build the data and replicate the findings of the main paper and appendix. Locals at the beginning of the do-file provide relevant directory links and indicators for which process to run. Users should create a folder called "Replication" which holds three further subfolders: "data", "figures", and "regs". Place all relevant data in "data", and replace the text in the local titled "fileloc" with the directory path to the "Replication" folder.

4 Software used

Stata/SE 13.1