Private Capital Research Institute

Data User Manual

Private & Confidential

** FOR APPROVED RESEARCHERS ONLY **

PRIOR TO OBTAINING ACCESS TO DATA, APPROVED DATA USERS WILL BE REQUIRED TO SIGN A DATA USER AGREEMENT.   ACCESS WILL BE LIMITED TO THE CONDITIONS OUTLINED IN THE DATA USER AGREEMENT.

Version 2.4

Table of Contents

What's new in PCRI database Version 2.4?

- Beginning in Spring 2016, the NORC migrated to a web-based portal, away from using thin client machines. This has enhanced researchers' experience in accessing the PCRI databases.

- For data security purposes, the following STATA commands are now disabled: list, browse, display, outsheet, and export

- Data Table Relationship Diagram

- **New!** Deal table which includes information about private capital transactions at the deal level

    o new variables include: *round, pe_equity, pe_debt, pe_value, total_value, tev.* See Appendix for more details.

- **New!** Data from Venture Intelligence, a leading source of information on private company financials, transactions (private equity, venture capital and M&A) and their valuations in India

- **New!** Data from StartUp Nation, a company that focuses on collecting data on Israeli private equity transactions and funds

- **New!** Improved industry codes: *industry_subtype* variable, which provides a more refined industry classification. See Appendix for detail.

- Updated Thomson Reuters, EMPEA, and unquote data through June 2016

- Data directly from more than 30 general partners

- Renaming of *nvca* variable to *industry* variable.

- *date* variable renamed to *deal_date, exit_date* in the Deal/Investment and Exit Tables, respectively.

- Investment Table now contains a direct link to the Exit Table via an *exit_id* variable

- Exit_type now includes an **assumed failure** type, which refers to buyout/growth equity and venture capital portfolio companies that have not received funding for ten years.

About the Private Capital Research Institute

The Private Capital Research Institute (PCRI) is a non-profit organization that works to further the understanding of the economic impact of private capital through independent academic studies.   PCRI's primary goal is to produce high quality academic research on the private capital industry based in large part on the comprehensive, centralized academic databases of private capital activity that the PCRI is building.

For questions regarding the data, please email *info@pcrinstitute.org* or visit our website at *www.pcrinstitute.org*.

Conditions of Data Access

a. Access to the Data Sets is limited solely to Requestor. Requestor may not provide, give, lend, or in any way make accessible the Data Sets to any third party, including, but not limited to, other institutions or researchers with which Requestor is affiliated, unless any such person has signed and submitted a Data User Agreement.

b. Requestor agrees to fulfill all pre-access Requirements

   i. Requestor and, if Requestor is affiliated with an institution, either the head of Requestor's department at such institution or the Dean of Requestor's institution, in both cases with institutional signature authority, must sign this Agreement.

   ii. Requestor must provide NORC with a pre-defined static or limited range IP Address from which Requestor will access data remotely, as access will only be permitted from this IP Address.

   iii. Requestor shall use Data Sets solely as stated in the Proposal. Data Sets may not be used for any other purposes whatsoever, including, but not limited to, consulting or other activities of Requestor and administrative, commercial, regulatory, law enforcement, judicial, or other purposes. Requestor shall not download substantial portions of data from the Data Sets.

c. In exchange for a lease fee payment by Requestor paid to NORC, NORC shall provide to Requestor a Thin Client Machine ("Thin Client"), which shall serve as Requestor's access node to the Data Sets. Requestor shall make no attempt to tamper with the configuration or security features of the Thin Client. Requestor shall use his or her best efforts to prevent unauthorized users from accessing the Thin Client. The Thin Client is the property of NORC and shall be returned to NORC at the expiration or termination of this Agreement.

d. Requestor shall not attempt to compromise the security of the computing environment by circumventing any access controls, sharing passwords or rights of access to the Data Sets or by capturing or sharing any images or information from the Data Sets. Without limiting the foregoing, Requestor shall not use screen capture software or devices, screen sharing software or devices or allow unauthorized users to view the Data Sets.

e. At PCRI's discretion, the names and addresses of any companies for which the Data Sets contain data may be replaced with a new variable name or address that does not identify the company. Requestor shall not attempt to identify providers of individual items of data to the Data Sets other than as disclosed by NORC and/or PCRI. Within 24 hours of the time when a Requestor becomes aware that any such information may have been identified, Requestor shall notify NORC and PCRI of the disclosure, inadvertent or otherwise. Requestor shall use his or her best efforts to destroy, retract or otherwise safeguard from further dissemination the materials containing such information.

f. All research output, such as statistical results and reports derived using Stata and SAS (or any other statistical analysis programs), must be reviewed by NORC and/or PCRI to ensure that the disclosure protection standards are met.

g. Requestor agrees to provide any work containing data derived from the Data Sets for review by NORC and PCRI at least thirty (30) days prior to the submission of such work for publication. Requestor shall not publish any work containing data derived from Data Sets without first deleting NORC's and/or PCRI's confidential and/or proprietary information as identified by NORC and/or PCRI during their review.

h. Requestor agrees to provide NORC and PCRI with copies of any work containing data derived from Data Sets within ninety (90) days of the publication date of such work.

i. Requestor agrees to credit PCRI and the NORC Data Enclave, as specified by NORC and PCRI, for any work containing data derived from Data Sets. In no event shall any such credit state or imply that NORC or PCRI endorse Requestor's research or research results.

Confidentiality

Except as authorized under this Agreement, Requestor shall not use or disclose any information obtained through the Data Sets or Data Enclave or any other proprietary or confidential information provided by PCRI or NORC and identified by them as such ("Confidential Information"). In no case shall Requestor use Confidential Information to identify by name or distinguishing characteristics any company of which Requestor learns about in the course of Requestor's research. Except as permitted and in connection with published research, Requestor shall keep all Confidential Information in confidence during and after the term of this Agreement unless any such Confidential Information (i) is or becomes generally publicly known and available, through no act or omission by the Requestor, or (ii) was rightfully in Requestor's possession free of any obligation of confidence at or prior to the time it was used or disclosed by Requestor. Requestor may disclose Confidential Information as and to the extent required by a valid order of a court or other governmental body, as otherwise required by law. Upon the request of NORC or PCRI, Requestor shall return to NORC and/or PCRI, as applicable, any Confidential Information in tangible form within ten (10) business days of such request.

Breaches

NORC shall promptly notify PCRI of any material breach of this Agreement by Requestor. NORC and PCRI may treat any breaches of this Agreement as violations of their policies and procedures on scientific integrity and misconduct. If NORC and PCRI deem Requestor to have breached this Agreement, each of them may:

a. Deny Requestor any further access to the Data Sets and the Data Enclave.

b. Withhold from Requestor undelivered data output from Requestor.

c. Report the violation to the appropriate authorities at Requestor's institution (and other authorities, as applicable and appropriate) and recommend sanctions be imposed.

d. Pursue all other remedies that may be available to under law or equity, including injunctive relief to prevent Requestor's use of any data derived from the Data Sets.

e. Terminate this Agreement.

Monitoring Usage

a. NORC and PCRI may take any reasonable steps to monitor Requestor's use of the Data Sets and access to the Data Enclave to ensure that Requestor complies with this Agreement. Such steps may include, without limitation, the following:

    i. A Data Enclave researcher assistant or member of PCRI may contact Requestor to confirm compliance with this Agreement.

    ii. NORC may maintain a full record of all computer-based interactions with the Data Sets at a keystroke level of detail and provide PCRI with such records.

    iii. NORC and PCRI may review all data output derived by Requestor from the Data Sets.

b. PCRI has the right to share the following information for the purpose of complying with any reporting or auditing obligations to its data providers:

    i. Requestor's name, address, e-mail address, telephone number, and name of the Requestor's department at his or her institution (if applicable); and

    ii. The name and address of Requestor's institution.
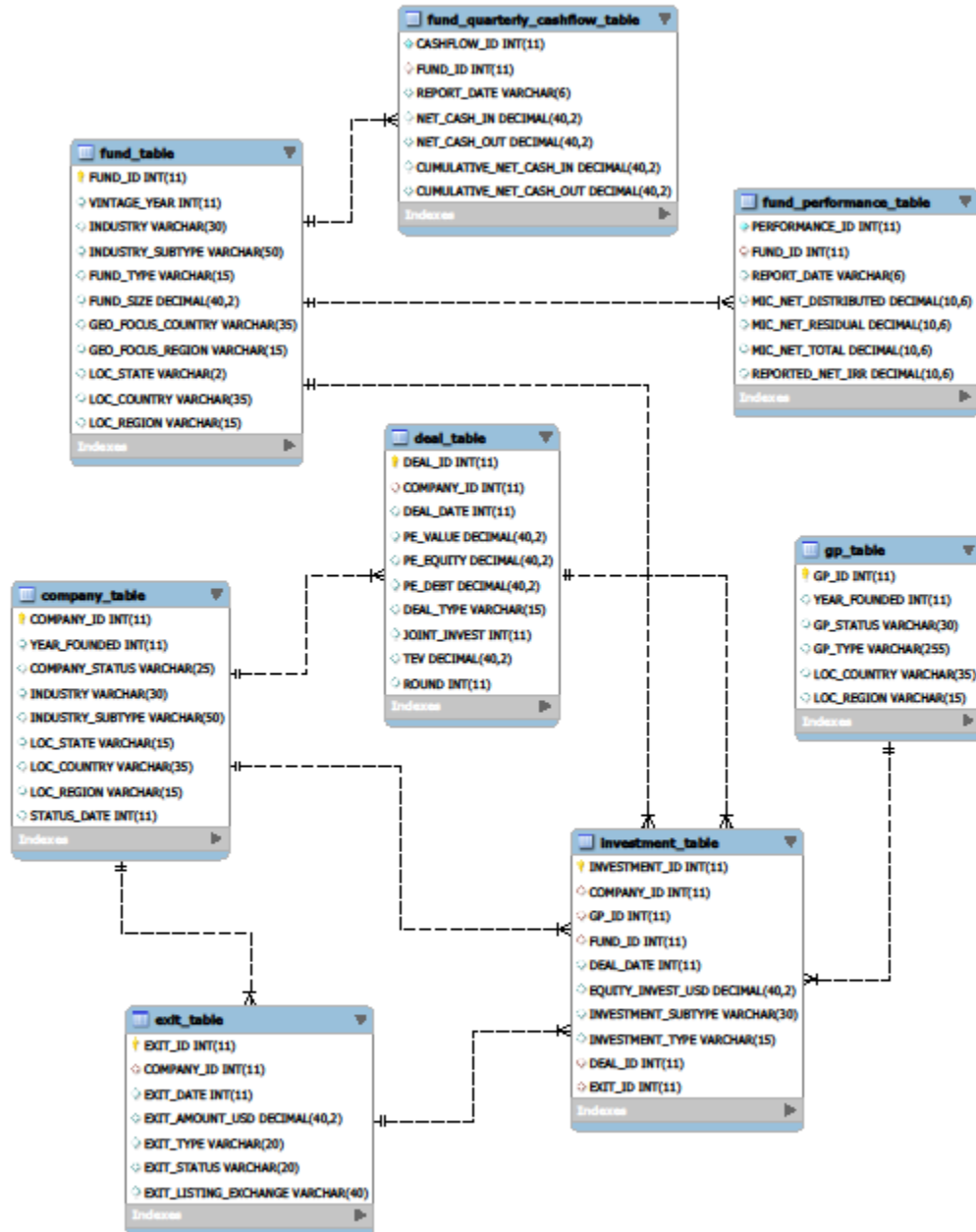
Table Relationship Diagram

Table Descriptions

Company Table

`\\enclave.local\pcri\Datasets\company.dta`

Information on portfolio companies involved in private capital transactions.

Variable Name	Definition
company_id	Unique identifier of a portfolio company. **Primary Key**.
year_founded	Self-reported year the portfolio company was founded
industry[1]	Industry code for the portfolio company based on the National Venture Capital Association 17 industry categories. See appendix for the industry codes. Labeled "Multi-sector" if multiple industry codes are determined.
industry_subtype	PCRI-determined classification into 74 subindustries. See Appendix for industry subtype codes.
company_status	Status of the company as of status date. See Appendix for definitions of status classifications.
status_date	As of date for company status. If unavailable, then the status date is current date.
loc_state	State of portfolio company's headquarters. See appendix for state codes. Missing if company is not headquartered in United States
loc_country	Country of portfolio company's headquarters. See appendix for country codes. Labeled "MULTINATIONAL" if company lists multiple countries
loc_region	Region of portfolio company's headquarters. See appendix for region codes. Labeled "MULTIGEOGRAPHY" if company lists multiple regions

---

[1] Industry classification is based on the 17 National Venture Capital Association (NVCA) categories. The PCRI creates mapping between SIC codes, GIC codes and business descriptions into the 17 NVCA industry categories.

Deal Table

`\\enclave.local\pcri\Datasets\deal.dta`

Information on private capital deals. A deal is defined as a set of investments into a company on a certain date.

Variable	Definition
deal_id	Unique identifier of a deal, which involves a company and date. **Primary Key**.
company_id	Unique identifier of the portfolio company. **Foreign Key**.
deal_date	Date of deal or investment round (in some instances, this is an estimated date); encoded as a Stata® date
round	Financing round
deal_type	Investments/deals are categorized into five main categories. See Appendix for investment types.
pe_value	Total amount of capital received by a company in millions of USD. *pe_value* is equal to the sum of *pe_equity* and *pe_debt.* In some instances, this is an estimate.
pe_equity	Amount of the deal that is equity in millions of USD. In some instances, this is an estimate.
pe_debt	Amount of the deal that is debt in millions of USD. In some instances, this is an estimate.
tev	Total Enterprise Value.
joint_invest	equals 1 if investment by multiple funds

Exit Table

`\\enclave.local\pcri\Datasets\exit.dta`

Information on Exit transactions of private capital investments: an exit is included if there is an associated private capital firm, date of transaction and an associated portfolio company.

Variable	Definition
exit_id	Unique identifier of an exit transaction. **Primary Key**.
company_id	Identifies the portfolio company in the exit transaction. **Foreign Key.**
exit_date	Date of exit transaction (possibly an estimated date); encoded as a Stata® date.
exit_amount_usd	Total size of exit consideration in millions of USD. In some cases, this is an estimated amount. Note: In syndicated deals, this is the total amount for all firms.
exit_type	Method of exit. See Appendix for type of exit.
exit_listing_exchange	Stock exchange where public shares are listed if IPO. See Appendix for list of exchanges. Variable set to missing if transaction was not an IPO.
exit_status	Label of how complete the exit is (realized, unrealized, or partially realized).

Fund Table

`\\enclave.local\pcri\Datasets\fund.dta`

Information on private capital funds.

Variable	Definition
fund_id	Unique identifier of a fund. Primary Key
vintage_year	Vintage year of fund or origination year of the fund. If Vintage Year is missing, we replace with the year of first investment of the fund.
industry[2]	Industry focus of the fund. Industry code for the portfolio company based on the National Venture Capital Association 17 industry categories. See appendix for the industry codes. Labeled "Multi-sector" if multiple industry codes are determined
industry_subtype	PCRI-determined classification into 74 subindustries. See Appendix for industry subtype codes.
fund_type	Fund's classification – see appendix for Fund_type
fund_size	Total committed amount of the fund in millions of US dollars.
geo_focus_country	Geographic focus of the fund. This variable is given or we infer from the portfolio companies within a fund. See Appendix for country codes.
geo_focus_region	Regional focus of the fund. This variable is given or we infer from the portfolio companies within a fund. See Appendix for region codes
loc_state	US State of fund's headquarters, if applicable. Missing if fund is not headquartered in United States. See Appendix for state codes.
loc_country	Country of fund's headquarters. See Appendix for country codes.

---

[2] Industry classification is based on the 17 National Venture Capital Association (NVCA) categories. The PCRI creates mapping between SIC codes, GIC codes and business descriptions into the 17 NVCA industry categories.

Fund Performance Table:

`\\enclave.local\pcri\Datasets\fund_performance.dta`

Information on private capital fund performance.

Variable	Definition
performance_id	Unique identifier of a fund in a specific quarter. **Primary Key**.
fund_id	Unique identifier of the fund. **Foreign Key.**
report_date	Date of performance report; stored as Stata® quarter
mic_net_distributed	Net distributed cash flow as a multiple of invested capital: Cumulative net cash out divided by cumulative cash in or $$\frac{\sum Cash\ out}{\sum Cash\ in}$$
mic_net_residual	Net residual as a multiple of invested capital: Total assets under management divided by cumulative net cash in or $$\frac{\sum Unrealized\ Assets}{\sum Cash\ in}$$
mic_net_total	Total value of net distributed and residual as a multiple of invested capital: cumulative net cash out plus total assets under management as a multiple of invested capital or $$\frac{\sum Cash\ out\ +\ Unrealized\ Assets}{\sum Cash\ in}$$
reported_net_irr	Reported IRR as of the report date, net of fees

Fund Quarterly Cashflow Table:

`\\enclave.local\pcri\Datasets\fund_quarterly_cashflow.dta`

Private capital fund cash flows.

Variable	Definition
cashflow_id	Unique identifier of a fund in a specific quarter. **Primary Key**.
fund_id	Identifier of the fund. **Foreign Key.**
report_date	Date of cash flow report; stored as Stata® quarter
net_cash_in[3]	Cash transferred into fund in a given quarter, net of fees, in thousands of US dollars.
net_cash_out[3]	Cash distributions from a fund in a given quarter, net of fees, in thousands of US dollars.
cumulative_net_cash_in	Sum of all cash transferred into a fund as of reported date, net of fees, in thousands of US dollars.
cumulative_net_cash_out	Sum of all cash distributions from a fund as of reported date, net of fees, in thousands of US dollars.

---

[3] The cash flows we receive typically represent the cash flow amounts for a particular fund for an *individual* limited partner. Consequently, we scale up an individual LP's cash flows to infer a total fund cash flow. Thus, in order to scale up, we need to have an LP's commitment as a percentage of the entire fund.

GP Table

`\\enclave.local\pcri\Datasets\gp.dta`

Information on the individual private capital firms.

Variable	Definition
gp_id	Unique identifier for private capital firm. **Primary Key**.
year_founded	Year the private capital firm was founded
gp_status	Investing status of firm. See Appendix for Status definitions.
gp_type	GP's classification – see Appendix for GP classifications.
loc_country	Country of firm's headquarters. See Appendix for country codes. Labeled "MULTINATIONAL" if firm lists multiple countries
loc_region	Region of firm's headquarters. See appendix for region codes. Labeled "MULTIGEOGRAPHY" if firm lists multiple regions

Investment Table

`\\enclave.local\pcri\Datasets\investment.dta`

Information on private capital investments.

Variable	Definition
investment_id	Unique identifier of an investment, which involves a company, private capital firm(s) and date. **Primary Key**.
company_id	Identifier of the portfolio company. **Foreign Key**.
fund_id	Identifier of the fund(s) in the investment. **Foreign Key**.
deal_date	Date of investment round (in some instances, this is an estimated date); encoded as a Stata® date
gp_id	Identifier of the investing firm(s). **Foreign Key**.
deal_id	Identifier of the deal to which an investment belongs, which involves a company and date. **Foreign Key.**
exit_id	Unique identifier of an exit transaction associated with an investment.
equity_invest_usd	Total capital invested by fund in millions USD. Includes common stock, convertible common stock, convertible preferred stock, warrants, options, debentures
investment_type	Investments/deals are categorized into five main categories. See Appendix for investment types.
investment_subtype	See Appendix for investment subtypes. 31 categories.

Keys

Keys provide the link between *two* tables and enable you to easily merge them. The merging process is explained in more detail in the next section. A Primary Key is a *unique* variable in a table. A Foreign Key is a variable that provides a link to another table. A Foreign Key is *not* necessarily unique in a table.

Fund Company Map

Variable	Definition
fund_id	Foreign key from Fund table
company_id	Foreign key from Company table

GP Company Map

Variable	Definition
gp_id	Foreign key from GP table
company_id	Foreign key from Company table

GP Fund Map

Variable	Definition
gp_id	Foreign key from GP table
fund_id	Foreign key from Fund table

Foreign Keys

Table Name	Primary Key	Foreign Key	Number of Rows
Company	company_id	none	138,408
Exit	exit_id	company_id, gp_id	55,544
Fund	fund_id	none	38,641
Fund_Performance	performance_id	fund_id	31,080
Fund_Quarterly_Cashflow	cashflow_id	fund_id	30,460
GP	gp_id	none	20,667
Investment	investment_id	company_id, fund_id, gp_id, exit_id deal_id	464,595
Deal	deal_id	company_id, gp_id	265,397

Fund_Company_Map	none	fund_id, company_id	180,449
GP_Company_Map	none	gp_id, company_id	263,718
GP_Fund_Map	none	gp_id, company_id	40,389

How to merge tables using the concept of keys

In order to merge tables, it is necessary to understand the concept of keys in a database. There are two types of keys: 1) primary and 2) foreign.

1) A <u>primary key</u> is a variable which *uniquely identifies observations* in a table.

investment_id	currency_id	fund_id	equity_invest_usd	…
1	EUR	84	25	…
2	USD	84	25	…
3	EUR	3	25	…
4	AUD	72	1.6	…
4	USD	72	1.6	…

Figure 1

In this case, investment_id is the primary key. Each of its values corresponds to exactly one unique row of data. Note: if the primary key appears more than once, the observation is a duplicate and should be removed using the duplicates drop command.

Currency_id, on the other hand, could *not* act as a primary key. This is because the value "USD" could refer to either one of the two lines highlighted above.

2) A <u>foreign key</u> is the *primary key of another table*. For instance, the fund_id variable in Figure 1 is actually the primary key for the following fund table:

fund_id	vintage_year	industry
1	2008	Multi-Sector
2	1997	Biotechnology
3	2003	Telecommunications

Figure 2

Foreign keys establish connections between tables by linking to unique rows of data. A table may have any number of foreign keys, but only one primary key. These concepts underpin the structure of <u>a relational database.</u>

Merges

The `merge` command in Stata allows the user to combine tables by *matching foreign keys to the respective tables in which they are primary keys*. There are many types of merges, but this guide will

cover those which will be used most frequently in conjunction with the PCRI database. For a more in-depth description, refer to the online Stata manual (http://www.stata.com/manuals13/dmerge.pdf).

1) The many-to-one merge (m:1)

   To perform a m:1 merge, we start by identifying the merge variable. This variable is the foreign key in one table and the primary key in the other.

   - The table containing the foreign key is called the <u>master dataset</u>
   - The table containing the primary key is called the <u>using dataset</u>

   A m:1 merge will link the *many* rows of data in the master dataset to *one* unique row in the using dataset. Doing this in Stata is straightforward: with the master dataset open, enter the merge command. The syntax is as follows:

   ```
   merge m:1 <merge_variable> using \\enclave.local\pcri\Datasets\<using_dataset>
   ```

   Input the `merge m:1` keyword, followed by the merge variable, followed by the file path of the using dataset. The file path for all PCRI tables will always begin as stated above.

   To illustrate this process, let us take a look at how to merge the Investment and Fund tables (Figures 1 and 2). The variable on which we merge is fund_id, since it is the foreign key in the Investment table and the primary key in the Fund table. Therefore, Investment.dta is the master dataset and Fund.dta is the using dataset. The image below illustrates the linking process:

Master dataset: Using dataset:

investment_id	…	fund_id
1	…	2
2	…	1
3	…	2
4	…	4
5	…	3
6	…	4

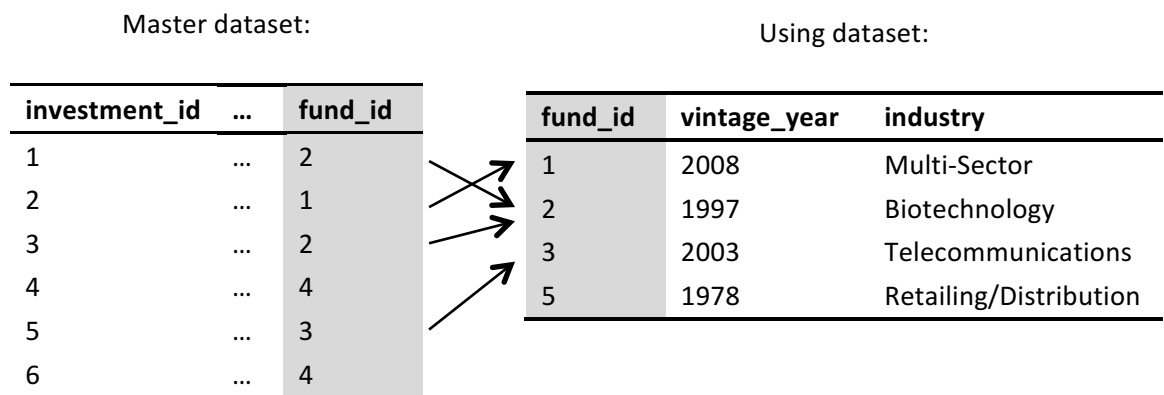| fund_id	vintage_year	industry
1	2008	Multi-Sector
2	1997	Biotechnology
3	2003	Telecommunications
5	1978	Retailing/Distribution

Figure 3

   The result of our merge is a wider and longer table of investments with vintage_year and industry information:

investment_id	…	fund_id	vintage_year	industry	_merge
1	…	2	1997	Biotechnology	3
2	…	1	2008	Multi-Sector	3
3	…	2	1997	Biotechnology	3
4	…	4	.		1
5	…	3	2003	Telecommunications	3
6	…	4	.		1
.	…	5	1978	Retailing/Distribution	2

Figure 4

Note that some data points are missing. This is because an instance of fund_id might appear in one table but not in the other. To keep track of this, Stata automatically adds an additional variable, called _merge, which can take on one of three values:

I.   _merge equals 3

The merge variable shows up in both datasets, yielding a complete row of information.

II.   _merge equals 2

The merge variable only shows up in the using dataset, yielding null values for the master dataset observations. Since these missing variables include the newly formed table's primary key, it is customary to drop these observations.

III.   _merge equals 1

The merge variable only shows up in the master dataset, yielding null values for the using dataset observations. It is optional whether or not to drop these lines.

Upon completing the merge, Stata displays a table containing the number of observations for each of these outcomes. Use this table to determine how effectively the tables have been merged. Ideally, we will find a large number of observations for which _merge equals 3. The _merge variable is typically dropped after these operations are completed (`drop _merge`).

In conclusion, the full Stata input for this scenario would be:

22

```
merge m:1 fund_id using \\enclave.local\pcri\Datasets\Fund

drop if _merge==2

drop _merge
```

2) The many-to-many merge (joinby)

There could be situations in which there aren't primary-foreign key pairs on which to merge. In these cases, a m:m merge is required. A m:m merge will link the *many* rows of data in the master dataset to *many* rows in the using dataset.

There is are very few scenarios within the PCRI database which call for such a merge. It is advisable to avoid performing this merge with large datasets, as this may cause runtime issues.

To perform a m:m merge, we start by identifying the merge variable. For instance, the purpose of merging the Investment and Exit tables would be to view all the investments and exits over the lifecycle of a company. Therefore, company_id is the merge variable. Note that it acts as the foreign key in *both* tables, meaning the master and using datasets are interchangeable. However, it is best practice to make the Investment.dta the master dataset. Once it is open, enter the following Stata command:

```
joinby company_id using \\enclave.local\pcri\Datasets\Exit, unmatched(both)
```

Input the `joinby` keyword, followed by the merge variable, followed by the file path of the using dataset. The resulting _merge variable has the same meaning as in the previous section, but with different implications:

I.   **_merge equals 3**

The merge variable shows up in both datasets, yielding a complete row of information.

II.   **_merge equals 2**

The merge variable only shows up in the using dataset, yielding null values for the master dataset observations. This probably means that the company investment occurred before the timeframe of the PCRI database or that the observations were not provided by PCRI vendors or GP's.

III.   **_merge equals 1**

The merge variable only shows up in the master dataset, yielding null values for the using dataset observations. This can mean one of three things:
1. Exit information is not known.
2. The company has not exited.
3. The company is defunct.

23

It is up to the researcher's discretion to determine which observations to keep.

Appendix

company_status	Description
public	Public company
private	Private company
subsidiary	Subsidiary company
defunct	Out of business
other	Government-owned, filed for bankruptcy

deal_type	Comprised of Types
buyout	Acquisition Finance, Acquisition for Expansion, Acquisition, Secondary Buyout, Corporate carveout, Industrial Holding Investment, LBO, Pending Acquisition, PIPE
vc	Early Stage, Seed, Start-up, VC Partnership
second	LP interest, Secondary Purchase
growth equity	Expansion/Growth, Later Stage
other	Bridge Loan, Direct, Mezzanine, Open Market Purchase, Other, Real Estate, Project, Turnaround, Debt Investment

exit_listing_exchange	Stock Exchange Name
AEX	Euronext Amsterdam, Amsterdam Stock Exchange
BIST	Istanbul Stock Exchange
BOSE	Botswana Stock Exchange
BSE	Bombay Stock Exchange
BSE-B	Boston Stock Exchange
BUX	Budapest Stock Exchange
CASE	Casablanca Stock Exchange
CSE	Copenhagen Stock Exchange
EBR	Euronext Brussels
FWB	Frankfurt Stock Exchange/ Neuer Markt
HKSE	Hong Kong Stock Exchange
ISE	Irish Stock Exchange (now part of LSE)
MLSE	Borso Italiano/ Nuovo Mercato/ Milan Stock Exchange

NISE	Nigerian Stock Exchange
NSE	National Stock Exchange of India
OMX	Nordic Stock Exchange Group
OSE	Oslo Bord/ Norwegian Stock Exchange
PAR	Euronext Paris/ Nouveau Marche/ Reglement Mensuel
SSE	Shanghai Stock Exchange
TYO	Tokyo Stock Exchange

exit_type	Definition
acquisition	Sale of company to investors
ipo	Initial Public Offering, or the first sale of stock by a private company to the public
defunct	includes bankruptcies and write-offs
assumed failure	a company that does not receive funding ten years after its last funding is assumed to have failed
other	dividends, income distribution (distribution of earnings to investors by issuing dividends) and stock distribution (distribution of earnings to investors by issuing stock/equity stake) and other distributions. Also includes restructures which are recapitalizations, refinancings, restructures, and dividend recapitalization (occurs when a company incurs new debt to pay a special dividend to private investors or shareholders. This usually involves a company owned by a private investment firm that can authorize a dividend recap as an alternative to selling its equity stake in the company).

fund_type	*Includes:*
buyout	Buyout, Multi-stage, Balanced Stage, Core, Energy, Generalist, Turnaround
vc	Early Stage, Seed Stage, Venture Capital
second	Secondary
growth equity	Growth, Later Stage
other	Fund of Funds, Hedge Fund, Infrastructure, Mezzanine, Microfinance, Other, Real Estate, Quoted, Pre-IPO, Value-Add, Co-investment/Annex, Microfinance, Opportunistic, Mutual Funds

geo_focus_country, geo_focus_region –   See loc_country and loc_region below

gp_status

seeking new investments

reduced investment activities

defunct

Inactive/unknown

gp_type	Includes:
buyout	Buyout, Generalist Private Equity
vc	Venture Capital
other	Mezzanine, Fund of Funds, Other Private Equity, Corporate

industry   (industry classification is based on National Venture Capital Association industry classifications)

biotechnology

business products and services

computer software

computers and peripherals

consumer products and services

electronics/instrumentation

financial services

healthcare services

industrial/energy

it services

media/entertainment

medical devices and equipment

multisector

networking and equipment

other

retailing/distribution

semiconductors

telecommunications

Industry_subtype		
administrative and accounting services	electronic office equipment	networking and communication
aerospace and defense	emergency services	oil and gas
agriculture, forestry and fishing	enterprise application software	peripherals
alternative energy		personal computers
analytical and scientific instrumentation	fintech	pharmaceuticals
artificial intelligence and machine learning	fixed line telecommunications	power supply
asset managers	food, drug and beverage retailers	private capital investors
automotive repair	government	programming and software services
Banks	healthcare	publishing
biotech for food and industrial applications	hedge fund	real estate
biotechnology		
broadcasting	hospitality	recreational and leisure products
business support services	hospitals, clinics and primary care	recreational facilities
clothing, accessories and personal products	hotel and lodging	restaurants
commodity chemicals	household goods and home construction	retailing services
computer graphics and digital imaging	industrial suppliers	security systems
computer services	insurance	semiconductors
construction and materials	internet	servers, workstations and terminals
consulting	long term and dependent care	social*
cpu and mainframes	machinery	specialty chemicals
cyber security	managed care	systems software
data services	media agencies	telecommunications equipment
diagnostic medical equipment and services	medical health products	therapeutic medical equipment
e-commerce	memory devices	transportation
educational services	mining	travel and leisure
electronic components	mobile telecommunications	utilities
		waste and disposal services

includes social, scientific, and public service investments

Investment_type	Comprised of Types
buyout	Acquisition Finance, Acquisition for Expansion, Acquisition, Secondary Buyout, Corporate carveout, Industrial Holding Investment, LBO, Pending Acquisition, PIPE
vc	Early Stage, Seed, Start-up, VC Partnership
second	LP interest, Secondary Purchase
growth equity	Expansion/Growth, Later Stage
other	Bridge Loan, Direct, Mezzanine, Open Market Purchase, Other, Real Estate, Project, Turnaround, Debt Investment

investment_subtype	
acquisition	pending acquisition
acquisition finance	private
acquisition for expansion	project
bridge loan	public-to-private
carve-out	real estate
control buyout	secondary buyout
direct	secondary purchase
distressed debt purchases	secondary-lbo
early stage	seed
expansion/growth	start-up
industrial holding investment	turnaround
lbo	vc partnership
lp interest	venture capital
later stage	
mezzanine	
open market purchase	
other	
pipe	

loc_country, geo_focus_country

afghanistan	gambia	nicaragua
albania	georgia	nigeria
algeria	germany	norway
angola	ghana	oman
argentina	gibraltar	pakistan
armenia	greece	palestine
australia	guatemala	panama
austria	guinea	papua new guinea
azerbaijan	guyana	paraguay
bahamas	honduras	peru
bahrain	hong kong	philippines
bangladesh	hungary	poland
barbados	iceland	portugal
belarus	india	qatar
belgium	indonesia	republic of congo
belize	iraq	republic of macedonia
bermuda	ireland	romania
bolivia	israel	russia
bosnia and herzegovina	italy	rwanda
botswana	jamaica	saudi arabia
bouvet island	japan	senegal
brazil	jordan	serbia
brunei darussalam	kazakhstan	sierra leone
bulgaria	kenya	singapore
burkina faso	kuwait	slovakia
burundi	kyrgyzstan	slovenia
cambodia	laos	solomon islands
cameroon	latvia	south africa
canada	lebanon	south korea
cayman islands	liberia	spain
central african republic	libya	sri lanka
channel islands	liechtenstein	sudan
chile	lithuania	swaziland
china	luxembourg	sweden
colombia	macao	switzerland
comoros	macedonia	syria
costa rica	madagascar	taiwan
cote d'ivoire	malawi	tajikistan

	malaysia	tanzania
croatia	maldives	thailand
cuba	mali	togo
cyprus	malta	trinidad and tobago
czech republic	marshall islands	tunisia
democratic republic of the congo	mauritania	turkey
denmark	mauritius	turkmenistan
dominica	mexico	uganda
dominican republic	moldova	ukraine
ecuador	monaco	united arab emirates
egypt	mongolia	united kingdom
el salvador	montserrat	united states
equatorial guinea	morocco	uruguay
estonia	mozambique	uzbekistan
ethiopia	multinational	venezuela
fiji	myanmar	vietnam
finland	namibia	zambia
france	netherlands	zimbabwe
gabon	new zealand	

loc_region, geo_focus_region

africa		
	algeria	mali
	angola	mauritania
	botswana	mauritius
	burkina faso	monaco
	burundi	morocco
	cameroon	mozambique
	central african republic	multinational
	chad	namibia
	comoros	nigeria
	cote d'ivoire	republic of congo
	democratic republic of the congo	rwanda
	djibouti	senegal
	equatorial guinea	sierra leone
	ethiopia	south africa
	gabon	sudan
	gambia	swaziland
	ghana	tanzania
	guinea	togo
	kenya	tunisia
	liberia	uganda
	libya	zambia
	madagascar	zimbabwe
	malawi	

antarctica	bouvet island

asia		
	afghanistan	maldives
	bangladesh	mongolia
	brunei darussalam	myanmar
	cambodia	pakistan
	china	philippines
	fiji	singapore
	hong kong	south korea
	india	sri lanka
	indonesia	taiwan
	japan	tajikistan
	kyrgyzstan	thailand
	laos	turkmenistan
	macao	uzbekistan

	malaysia	vietnam
eurasia	armenia	
	azerbaijan	
	cyprus	
	kazakhstan	
	russia	
europe	albania	liechtenstein
	austria	lithuania
	belarus	luxembourg
	belgium	macedonia
	bosnia and herzegovina	malta
	bulgaria	moldova
	channel islands	multinational
	croatia	netherlands
	czech republic	norway
	denmark	poland
	estonia	portugal
	finland	republic of macedonia
	france	romania
	georgia	serbia
	germany	slovakia
	gibraltar	slovenia
	greece	spain
	hungary	sweden
	iceland	switzerland
	ireland	ukraine
	italy	united kingdom
	latvia	
middle east	bahrain	oman
	egypt	palestine
	iraq	qatar
	israel	saudi arabia
	jordan	syria
	kuwait	turkey
	lebanon	united arab emirates
multigeography	multinational	

north america	bahamas	el salvador
	barbados	guatemala
	belize	honduras
	bermuda	jamaica
	canada	mexico
	cayman islands	montserrat
	costa rica	nicaragua
	cuba	panama
	dominica	trinidad and tobago
	dominican republic	united states
oceania	australia	
	marshall islands	
	new zealand	
	papua new guinea	
south america	argentina	guyana
	bolivia	paraguay
	brazil	peru
	chile	uruguay
	colombia	venezuela
	ecuador	

loc_state

AK - Alaska	KY - Kentucky	OH - Ohio
AL - Alabama	LA - Louisiana	OK - Oklahoma
AR - Arkansas	MA - Massachusetts	OR - Oregon
AZ - Arizona	MD - Maryland	PA - Pennsylvania
CA - California	ME - Maine	PR - Puerto Rico
CO - Colorado	MI - Michigan	RI - Rhode Island
CT - Connecticut	MN - Minnesota	SC - South Carolina
DC - District of Columbia	MO - Missouri	SD - South Dakota
DE - Delaware	MS - Mississippi	TN - Tennessee
FL - Florida	MT - Montana	TX - Texas
GA - Georgia	NC - North Carolina	UT - Utah
GU - Guam	ND - North Dakota	VA - Virginia
HI - Hawaii	NE - Nebraska	VT - Vermont
IA - Iowa	NH - New Hampshire	WA - Washington
ID - Idaho	NJ - New Jersey	WI - Wisconsin
IL - Illinois	NM - New Mexico	WV - West Virginia
IN - Indiana	NV - Nevada	WY - Wyoming
KS - Kansas	NY - New York	

PCRI

Sample Data

Company

company_id	year_founded	company_status	status_date	industry	industry_subtype	loc_state	loc_country	loc_region
1	2005	private	132051	consumer products and services	household goods and home construction	MA	united states	north america
2	1997	public	132051	industrial/energy	machinery		india	asia
3	2001	subsidiary	90543	it services	e-commerce	CA	united states	north america

Exit

exit_id	company_id	exit_date	exit_amount_usd	exit_type	exit_listing_exchange	exit_status
1	12345	76253	23.7	buy-back		Realized
2	23456	132051	731	acquisition		Realized
3	34567	1321	17	ipo	NASDAQ	Partially Realized

PCRI

Fund

fund_id	vintage_year	industry	industry_subtype	fund_size	fund_type	geo_focus_country	geo_focus_region	…
1	2008	multisector		100	buyout	multinational	europe	…
2	1997	biotechnology	pharmaceuticals	2500	vc	united states	north america	…
3	2003	telecommunications	Mobile telecommunications	160	buyout	united states	north america	…

…	loc_state	loc_country	loc_region
…		multinational	europe
…	CA	united states	north america
…	MA	united states	north america

Fund Performance

performance_id	fund_id	report_date	mic_net_distributed	mic_net_residual	mic_net_total	reported_net_irr
1	13	2010Q3	0.91	0.77	1.688	0.141
2	13	2010Q4	0.92	0.69	1.617	0.149
3	57	2007Q1	1.78	0.401	2.23	0.212

Fund Quarterly Cashflow

cashflow_id	fund_id	report_date	net_cash_in	net_cash_out	cumulative_net_cash_in	cumulative_net_cash_out
543	345	2010Q2	5.9	0.1	118.46	3.56
544	345	2010Q3	42.6	1.6	163.4	5.81
545	345	2011Q2	13.8	55	191.76	61.2

GP

gp_id	year_founded	gp_status	gp_type	loc_country	loc_region
1	2005	seeking new investments	other	united states	north america
2	1997	inactive/unknown	vc	united kingdom	europe
3	2001	defunct	buyout	brazil	south america

Investment

investment_id	company_id	gp_id	fund_id	deal_date	deal_id	exit_id	equity_invest_usd	investment_type
1	708	5	84	132051	4	221	150	buyout
2	5	41	145	132050	108	16	200	vc
3	15	132	3	102344	27	1053	175	Other

Deal

deal_id	company_id	deal_date	deal_type	pe_value	pe_equity	pe_debt	tev	joint_invest
4	708	42301	buyout	1000	950	50	3000	0
108	5	88812	vc	80	80	0	400	1
27	15	126751	growth equity	200	175	25	800	0